PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : C12N 15/12, C12Q 1/68, C07K 14/47, 16/18, A01K 67/027, A61K 38/17, G01N 33/68	A2	(11) International Publication Number: WO 00/70047 (43) International Publication Date: 23 November 2000 (23.11.00)

(21) International Application Number: PCT/US00/13299

(22) International Filing Date: 12 May 2000 (12.05.00)

(30) Priority Data:

09/311,894	14 May 1999 (14.05.99)	US
09/311,940	14 May 1999 (14.05.99)	US
09/311,937	14 May 1999 (14.05.99)	US

(63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Applications

US	09,311,894 (CIP)
Filed on	14 May 1999 (14.05.99)
US	09,311,940 (CIP)
Filed on	14 May 1999 (14.05.99)
US	09,311,937 (CIP)
Filed on	14 May 1999 (14.05.99)

(71) Applicant (*for all designated States except US*): INCYTE GENOMICS, INC. [US/US]; 3160 Porter Drive, Palo Alto, CA 94304 (US).

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): YUE, Henry [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US). TANG, Y., Tom [CN/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). LAL, Preeti [IN/US]; 2382 Lass Drive, Santa Clara, CA 95054 (US). REDDY, Roopa [IN/US]; 1233 W. McKinley Drive, Sunnyvale, CA 94086 (US). BATRA, Sajeev [US/US]; 555 El Camino Real #709, Sunnyvale, CA 94087 (US). BAUGHN, Mariah, R. [US/US]; 14244 Santiago Road, San Leandro, CA 94577 (US). YANG, Junming [CN/US]; 7125 Bark Lane, San Jose, CA 95129 (US). AZIMZAI, Yalda [US/US]; 2045 Rock Springs Drive, Hayward, CA 94545 (US). LU, Dyung, Aina, M. [US/US]; 55 Park Belmont Place, San Jose, CA 95136 (US). AU-YOUNG, Janice [US/US]; 233 Golden Eagle Lane, Brisbane, CA 94005 (US). SHIH, Leo, L. [US/US]; Apartment B., 1081 Tanland Drive, Palo Alto, CA 94303 (US).

(74) Agents: HAMLET-COX, Diana et al.; Incyte Genomics, 3160 Porter Drive, Palo Alto, CA 94304 (US).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published

Without international search report and to be republished upon receipt of that report.

(54) Title: FULL-LENGTH MOLECULES EXPRESSED IN HUMAN TISSUES

(57) Abstract

The invention provides full-length molecules expressed in human tissues (FLEXHT) and polynucleotides which identify and encode FLEXHT. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also provides methods for analyzing human gene expression, genetic linkage, and genetic variability, and diagnosing, treating, or preventing disorders associated with expression of FLEXHT.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania
AM	Armenia
AT	Austria
AU	Australia
AZ	Azerbaijan
BA	Bosnia and Herzegovina
BB	Barbados
BE	Belgium
BF	Burkina Faso
BG	Bulgaria
BJ	Benin
BR	Brazil
BY	Belarus
CA	Canada
CF	Central African Republic
CG	Congo
CH	Switzerland
CI	Côte d'Ivoire
CM	Cameroon
CN	China
CU	Cuba
CZ	Czech Republic
DE	Germany
DK	Denmark
EE	Estonia
ES	Spain
FI	Finland
FR	France
GA	Gabon
GB	United Kingdom
GE	Georgia
GH	Ghana
GN	Guinea
GR	Greece
HU	Hungary
IE	Ireland
IL	Israel
IS	Iceland
IT	Italy
JP	Japan
KE	Kenya
KG	Kyrgyzstan
KP	Democratic People's Republic of Korea
KR	Republic of Korea
KZ	Kazakstan
LC	Saint Lucia
LI	Liechtenstein
LK	Sri Lanka
LR	Liberia
LS	Lesotho
LT	Lithuania
LU	Luxembourg
LV	Latvia
MC	Monaco
MD	Republic of Moldova
MG	Madagascar
MK	The former Yugoslav Republic of Macedonia
ML	Mali
MN	Mongolia
MR	Mauritania
MW	Malawi
MX	Mexico
NE	Niger
NL	Netherlands
NO	Norway
NZ	New Zealand
PL	Poland
PT	Portugal
RO	Romania
RU	Russian Federation
SD	Sudan
SE	Sweden
SG	Singapore
SI	Slovenia
SK	Slovakia
SN	Senegal
SZ	Swaziland
TD	Chad
TG	Togo
TJ	Tajikistan
TM	Turkmenistan
TR	Turkey
TT	Trinidad and Tobago
UA	Ukraine
UG	Uganda
US	United States of America
UZ	Uzbekistan
VN	Viet Nam
YU	Yugoslavia
ZW	Zimbabwe

FULL-LENGTH MOLECULES EXPRESSED IN HUMAN TISSUES

TECHNICAL FIELD

This invention relates to nucleic acid and amino acid sequences of full-length molecules expressed in human tissues and to the use of these sequences in the analysis of human gene expression, genetic linkage, and genetic variability, and in the diagnosis, treatment, and prevention of developmental, cell proliferative, and immunological disorders.

BACKGROUND OF THE INVENTION

It is estimated that only 2% of mammalian DNA encodes proteins, and only a small fraction of the genes that encode proteins is actually expressed in a particular cell at any time. The various types of cells in a multicellular organism differ dramatically both in structure and function, and the identity of a particular cell is conferred by its unique pattern of gene expression. In addition, different cell types express overlapping but distinctive sets of genes throughout development. Cell growth and proliferation, cell differentiation, the immune response, apoptosis, and other processes that contribute to organism development and survival are governed by regulation of gene expression. Appropriate gene regulation also ensures that cells function efficiently by expressing only those genes whose functions are required at a given time. Factors that influence gene expression include extracellular signals that mediate cell-cell communication and coordinate the activities of different cell types. Gene expression is regulated at the level of DNA and RNA transcription, and at the level of mRNA translation.

Aberrant expression or mutations in genes and their products may cause, or increase susceptibility to, a variety of human diseases such as cancer and other cell proliferative disorders. The identification of these genes and their products is the basis of an ever-expanding effort to finding markers for early detection of diseases and targets for their prevention and treatment. For example, cancer represents a type of cell proliferative disorder that affects nearly every tissue in the body. The development of cancer, or oncogenesis, is often correlated with the conversion of a normal gene into a cancer-causing gene, or oncogene, through abnormal expression or mutation. Oncoproteins, the products of oncogenes, include a variety of molecules that influence cell proliferation, such as growth factors, growth factor receptors, intracellular signal transducers, nuclear transcription factors, and cell-cycle control proteins. In contrast, tumor-suppressor genes are involved in inhibiting cell proliferation. Mutations which reduce or abrogate the function of tumor-suppressor genes result in aberrant cell proliferation and cancer. Thus a wide variety of genes and their products have been found that are associated with cell proliferative disorders such as cancer, but many more may exist that are yet to be discovered.

DNA-based arrays can provide an efficient, high-throughput method to examine gene expression and genetic variability. For example, SNPs, or single nucleotide polymorphisms, are the most common type of human genetic variation. DNA-based arrays can dramatically accelerate the discovery of SNPs in hundreds and even thousands of genes. Likewise, such arrays can be used for SNP genotyping in which DNA samples from individuals or populations are assayed for the presence of selected SNPs. These approaches will ultimately lead to the systematic identification of all genetic variations in the human genome and the correlation of certain genetic variations with disease susceptibility, responsiveness to drug treatments, and other medically relevant information. (See, for example, Wang, D.G. et al. (1998) Science 280:1077-1082.)

DNA-based array technology is especially important for the rapid analysis of global gene expression patterns. For example, genetic predisposition, disease, or therapeutic treatment may directly or indirectly affect the expression of a large number of genes in a given tissue. In this case, it is useful to develop a profile, or transcript image, of all the genes that are expressed and the levels at which they are expressed in that particular tissue. A profile generated from an individual or population affected with a certain disease or undergoing a particular therapy may be compared with a profile likewise generated from a control individual or population. Such analysis does not require knowledge of gene function, as the expression profiles can subjected to mathematical analyses which simply treat each gene as a marker. Furthermore, gene expression profiles may help dissect biological pathways by identifying all the genes expressed, for example, at a certain developmental stage, in a particular tissue, or in response to disease or treatment. (See, for example, Lander, E.S. et al. (1996) Science 274:536-539.)

The discovery of new full-length molecules expressed in human tissues and the polynucleotides encoding them satisfies a need in the art by providing new compositions which are useful in the analysis of human gene expression, genetic linkage, and genetic variability, and in the diagnosis, prevention, and treatment of developmental, cell proliferative, and immunological disorders.

SUMMARY OF THE INVENTION

The invention features purified polypeptides, full-length molecules expressed in human tissues, referred to collectively as "FLEXHT" and individually as "FLEXHT-1," "FLEXHT-2," "FLEXHT-3," "FLEXHT-4," "FLEXHT-5," "FLEXHT-6," "FLEXHT-7," "FLEXHT-8," "FLEXHT-9," "FLEXHT-10," "FLEXHT-11," "FLEXHT-12," "FLEXHT-13," "FLEXHT-14," "FLEXHT-15," "FLEXHT-16," "FLEXHT-17," "FLEXHT-18," "FLEXHT-19," "FLEXHT-20," "FLEXHT-21," "FLEXHT-22," "FLEXHT-23," "FLEXHT-24," "FLEXHT-25," "FLEXHT-26," "FLEXHT-27," "FLEXHT-28," "FLEXHT-29," "FLEXHT-30," "FLEXHT-31," "FLEXHT-32," "FLEXHT-33," "FLEXHT-34," "FLEXHT-35," "FLEXHT-36," "FLEXHT-37," "FLEXHT-38," "FLEXHT-39,"

"FLEXHT-40," "FLEXHT-41," "FLEXHT-42," "FLEXHT-43," "FLEXHT-44," "FLEXHT-45," "FLEXHT-46," "FLEXHT-47," "FLEXHT-48," "FLEXHT-49," "FLEXHT-50," "FLEXHT-51," "FLEXHT-52," "FLEXHT-53," "FLEXHT-54," and "FLEXHT-55." In one aspect, the invention provides an isolated polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-55, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-55, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-55, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-55. In one alternative, the invention provides an isolated polypeptide comprising the amino acid sequence of SEQ ID NO:1-55.

The invention further provides an isolated polynucleotide encoding a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-55, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-55, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-55, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-55. In one alternative, the polynucleotide encodes a polypeptide selected from the group consisting of SEQ ID NO:1-55. In another alternative, the polynucleotide is selected from the group consisting of SEQ ID NO:56-110.

Additionally, the invention provides a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-55, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-55, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-55, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-55. In one alternative, the invention provides a cell transformed with the recombinant polynucleotide. In another alternative, the invention provides a transgenic organism comprising the recombinant polynucleotide.

The invention also provides a method for producing a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-55, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-55, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO

NO:1-55, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-55. The method comprises a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding the polypeptide, and b) recovering the polypeptide so expressed.

Additionally, the invention provides an isolated antibody which specifically binds to a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-55, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-55, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-55, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-55.

The invention further provides an isolated polynucleotide comprising a polynucleotide sequence selected from the group consisting of a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:56-110, b) a naturally occurring polynucleotide sequence having at least 90% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:56-110, c) a polynucleotide sequence complementary to a), d) a polynucleotide sequence complementary to b), and e) an RNA equivalent of a)-d). In one alternative, the polynucleotide comprises at least 60 contiguous nucleotides.

Additionally, the invention provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide comprising a polynucleotide sequence selected from the group consisting of a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:56-110, b) a naturally occurring polynucleotide sequence having at least 90% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:56-110, c) a polynucleotide sequence complementary to a), d) a polynucleotide sequence complementary to b), and e) an RNA equivalent of a)-d). The method comprises a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and b) detecting the presence or absence of said hybridization complex, and optionally, if present, the amount thereof. In one alternative, the probe comprises at least 60 contiguous nucleotides.

The invention further provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide comprising a polynucleotide sequence

selected from the group consisting of a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:56-110, b) a naturally occurring polynucleotide sequence having at least 90% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:56-110, c) a polynucleotide sequence complementary to a), d) a polynucleotide sequence complementary to b), and e) an RNA equivalent of a)-d). The method comprises a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.

The invention further provides a pharmaceutical composition comprising an effective amount of a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-55, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-55, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-55, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-55, and a pharmaceutically acceptable excipient. In one embodiment, the pharmaceutical composition comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-55. The invention additionally provides a method of treating a disease or condition associated with decreased expression of functional FLEXHT, comprising administering to a patient in need of such treatment the pharmaceutical composition.

The invention also provides a method for screening a compound for effectiveness as an agonist of a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-55, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-55, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-55, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-55. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting agonist activity in the sample. In one alternative, the invention provides a pharmaceutical composition comprising an agonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of treating a disease or condition associated with decreased expression of functional FLEXHT, comprising administering to a patient in need of such treatment the pharmaceutical composition.

Additionally, the invention provides a method for screening a compound for effectiveness as an antagonist of a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-55, b) a naturally

occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-55, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-55, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-55. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting antagonist activity in the sample. In one alternative, the invention provides a pharmaceutical composition comprising an antagonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of treating a disease or condition associated with overexpression of functional FLEXHT, comprising administering to a patient in need of such treatment the pharmaceutical composition.

The invention further provides a method of screening for a compound that specifically binds to a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-55, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-55, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-55, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-55. The method comprises a) combining the polypeptide with at least one test compound under suitable conditions, and b) detecting binding of the polypeptide to the test compound, thereby identifying a compound that specifically binds to the polypeptide.

The invention further provides a method of screening for a compound that modulates the activity of a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-55, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-55, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-55, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-55. The method comprises a) combining the polypeptide with at least one test compound under conditions permissive for the activity of the polypeptide, b) assessing the activity of the polypeptide in the presence of the test compound, and c) comparing the activity of the polypeptide in the presence of the test compound with the activity of the polypeptide in the absence of the test compound, wherein a change in the activity of the polypeptide in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide.

The invention further provides a method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a

sequence selected from the group consisting of SEQ ID NO:56-110, the method comprising a) exposing a sample comprising the target polynucleotide to a compound, and b) detecting altered expression of the target polynucleotide.

The invention further provides an isolated polynucleotide comprising at least 20 contiguous nucleotides of a polynucleotide comprising a polynucleotide sequence selected from the group consisting of a) a polynucleotide sequence which is complementary to a polynucleotide sequence selected from the group consisting of SEQ ID NO:56-110, and b) an RNA equivalent of a). The invention also provides a composition comprising at least one of these isolated polynucleotides and a detectable label, said composition being useful for the detection of altered expression of human FLEXHT.

The invention further provides a microarray wherein at least one element of the microarray is an isolated polynucleotide comprising at least 20 contiguous nucleotides of a polynucleotide comprising a polynucleotide sequence selected from the group consisting of a) a polynucleotide sequence which is complementary to a polynucleotide sequence selected from the group consisting of SEQ ID NO:56-110, and b) an RNA equivalent of a). The invention also provides a method of using the microarray for generating a transcript image of a sample which contains polynucleotides. The method comprises a) labeling the polynucleotides of the sample, b) contacting the elements of the microarray with the labeled polynucleotides of the sample under conditions suitable for the formation of a hybridization complex, and c) quantifying the expression of the polynucleotides in the sample.

BRIEF DESCRIPTION OF THE TABLES

Table 1 shows polypeptide and nucleotide sequence identification numbers (SEQ ID NOs), clone identification numbers (clone IDs), cDNA libraries, and cDNA fragments used to assemble full-length sequences encoding FLEXHT.

Table 2 shows features of each polypeptide sequence, including protein phosphorylation and glycosylation sites.

Table 3 shows selected fragments of each nucleic acid sequence; the tissue-specific expression patterns of each nucleic acid sequence as determined by northern analysis; diseases, disorders, or conditions associated with these tissues; and the vector into which each cDNA was cloned.

Table 4 describes the tissues used to construct the cDNA libraries from which cDNA clones encoding FLEXHT were isolated.

Table 5 shows features of selected polypeptide sequences, including the identity of the polypeptide, potential motifs, homologous sequences, and methods, algorithms, and searchable databases used for analysis of FLEXHT.

Table 6 shows the tools, programs, and algorithms used to analyze the polynucleotides and polypeptides of the invention, along with applicable descriptions, references, and threshold parameters.

DESCRIPTION OF THE INVENTION

Before the present proteins, nucleotide sequences, and methods are described, it is understood that this invention is not limited to the particular machines, materials and methods described, as these may vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention which will be limited only by the appended claims.

It must be noted that as used herein and in the appended claims, the singular forms "a," "an," and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, a reference to "a host cell" includes a plurality of such host cells, and a reference to "an antibody" is a reference to one or more antibodies and equivalents thereof known to those skilled in the art, and so forth.

Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any machines, materials, and methods similar or equivalent to those described herein can be used to practice or test the present invention, the preferred machines, materials and methods are now described. All publications mentioned herein are cited for the purpose of describing and disclosing the cell lines, protocols, reagents and vectors which are reported in the publications and which might be used in connection with the invention. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

DEFINITIONS

"FLEXHT" refers to the amino acid sequences of substantially purified FLEXHT obtained from any species, particularly a mammalian species, including bovine, ovine, porcine, murine, equine, and human, and from any source, whether natural, synthetic, semi-synthetic, or recombinant.

The term "agonist" refers to a molecule which intensifies or mimics the biological activity of FLEXHT. Agonists may include proteins, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of FLEXHT either by directly interacting with FLEXHT or by acting on components of the biological pathway in which FLEXHT participates.

An "allelic variant" is an alternative form of the gene encoding FLEXHT. Allelic variants may result from at least one mutation in the nucleic acid sequence and may result in altered mRNAs or in polypeptides whose structure or function may or may not be altered. A gene may have none, one, or many allelic variants of its naturally occurring form. Common mutational changes which give rise to allelic variants are generally ascribed to natural deletions, additions, or substitutions of nucleotides. Each of these types of changes may occur alone, or in combination with the others, one or more times in

a given sequence.

"Altered" nucleic acid sequences encoding FLEXHT include those sequences with deletions, insertions, or substitutions of different nucleotides, resulting in a polypeptide the same as FLEXHT or a polypeptide with at least one functional characteristic of FLEXHT. Included within this definition are polymorphisms which may or may not be readily detectable using a particular oligonucleotide probe of the polynucleotide encoding FLEXHT, and improper or unexpected hybridization to allelic variants, with a locus other than the normal chromosomal locus for the polynucleotide sequence encoding FLEXHT. The encoded protein may also be "altered," and may contain deletions, insertions, or substitutions of amino acid residues which produce a silent change and result in a functionally equivalent FLEXHT. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues, as long as the biological or immunological activity of FLEXHT is retained. For example, negatively charged amino acids may include aspartic acid and glutamic acid, and positively charged amino acids may include lysine and arginine. Amino acids with uncharged polar side chains having similar hydrophilicity values may include: asparagine and glutamine; and serine and threonine. Amino acids with uncharged side chains having similar hydrophilicity values may include: leucine, isoleucine, and valine; glycine and alanine; and phenylalanine and tyrosine.

The terms "amino acid" and "amino acid sequence" refer to an oligopeptide, peptide, polypeptide, or protein sequence, or a fragment of any of these, and to naturally occurring or synthetic molecules. Where "amino acid sequence" is recited to refer to a sequence of a naturally occurring protein molecule, "amino acid sequence" and like terms are not meant to limit the amino acid sequence to the complete native amino acid sequence associated with the recited protein molecule.

"Amplification" relates to the production of additional copies of a nucleic acid sequence. Amplification is generally carried out using polymerase chain reaction (PCR) technologies well known in the art.

The term "antagonist" refers to a molecule which inhibits or attenuates the biological activity of FLEXHT. Antagonists may include proteins such as antibodies, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of FLEXHT either by directly interacting with FLEXHT or by acting on components of the biological pathway in which FLEXHT participates.

The term "antibody" refers to intact immunoglobulin molecules as well as to fragments thereof, such as Fab, $F(ab')_2$, and Fv fragments, which are capable of binding an epitopic determinant. Antibodies that bind FLEXHT polypeptides can be prepared using intact polypeptides or using fragments containing small peptides of interest as the immunizing antigen. The polypeptide or oligopeptide used to immunize an animal (e.g., a mouse, a rat, or a rabbit) can be derived from the

translation of RNA, or synthesized chemically, and can be conjugated to a carrier protein if desired. Commonly used carriers that are chemically coupled to peptides include bovine serum albumin, thyroglobulin, and keyhole limpet hemocyanin (KLH). The coupled peptide is then used to immunize the animal.

The term "antigenic determinant" refers to that region of a molecule (i.e., an epitope) that makes contact with a particular antibody. When a protein or a fragment of a protein is used to immunize a host animal, numerous regions of the protein may induce the production of antibodies which bind specifically to antigenic determinants (particular regions or three-dimensional structures on the protein). An antigenic determinant may compete with the intact antigen (i.e., the immunogen used to elicit the immune response) for binding to an antibody.

The term "antisense" refers to any composition capable of base-pairing with the "sense" (coding) strand of a specific nucleic acid sequence. Antisense compositions may include DNA; RNA; peptide nucleic acid (PNA); oligonucleotides having modified backbone linkages such as phosphorothioates, methylphosphonates, or benzylphosphonates; oligonucleotides having modified sugar groups such as 2'-methoxyethyl sugars or 2'-methoxyethoxy sugars; or oligonucleotides having modified bases such as 5-methyl cytosine, 2'-deoxyuracil, or 7-deaza-2'-deoxyguanosine. Antisense molecules may be produced by any method including chemical synthesis or transcription. Once introduced into a cell, the complementary antisense molecule base-pairs with a naturally occurring nucleic acid sequence produced by the cell to form duplexes which block either transcription or translation. The designation "negative" or "minus" can refer to the antisense strand, and the designation "positive" or "plus" can refer to the sense strand of a reference DNA molecule.

The term "biologically active" refers to a protein having structural, regulatory, or biochemical functions of a naturally occurring molecule. Likewise, "immunologically active" or "immunogenic" refers to the capability of the natural, recombinant, or synthetic FLEXHT, or of any oligopeptide thereof, to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

"Complementary" describes the relationship between two single-stranded nucleic acid sequences that anneal by base-pairing. For example, 5'-AGT-3' pairs with its complement, 3'-TCA-5'.

A "composition comprising a given polynucleotide sequence" and a "composition comprising a given amino acid sequence" refer broadly to any composition containing the given polynucleotide or amino acid sequence. The composition may comprise a dry formulation or an aqueous solution. Compositions comprising polynucleotide sequences encoding FLEXHT or fragments of FLEXHT may be employed as hybridization probes. The probes may be stored in freeze-dried form and may be

associated with a stabilizing agent such as a carbohydrate. In hybridizations, the probe may be deployed in an aqueous solution containing salts (e.g., NaCl), detergents (e.g., sodium dodecyl sulfate; SDS), and other components (e.g., Denhardt's solution, dry milk, salmon sperm DNA, etc.).

"Consensus sequence" refers to a nucleic acid sequence which has been subjected to repeated DNA sequence analysis to resolve uncalled bases, extended using the XL-PCR kit (Perkin-Elmer, Norwalk CT) in the 5' and/or the 3' direction, and resequenced, or which has been assembled from one or more overlapping cDNA, EST, or genomic DNA fragments using a computer program for fragment assembly, such as the GELVIEW fragment assembly system (GCG, Madison WI) or Phrap (University of Washington, Seattle WA). Some sequences have been both extended and assembled to produce the consensus sequence.

"Conservative amino acid substitutions" are those substitutions that are predicted to least interfere with the properties of the original protein, i.e., the structure and especially the function of the protein is conserved and not significantly changed by such substitutions. The table below shows amino acids which may be substituted for an original amino acid in a protein and which are regarded as conservative amino acid substitutions.

Original Residue	Conservative Substitution
Ala	Gly, Ser
Arg	His, Lys
Asn	Asp, Gln, His
Asp	Asn, Glu
Cys	Ala, Ser
Gln	Asn, Glu, His
Glu	Asp, Gln, His
Gly	Ala
His	Asn, Arg, Gln, Glu
Ile	Leu, Val
Leu	Ile, Val
Lys	Arg, Gln, Glu
Met	Leu, Ile
Phe	His, Met, Leu, Trp, Tyr
Ser	Cys, Thr
Thr	Ser, Val
Trp	Phe, Tyr
Tyr	His, Phe, Trp
Val	Ile, Leu, Thr

Conservative amino acid substitutions generally maintain (a) the structure of the polypeptide backbone in the area of the substitution, for example, as a beta sheet or alpha helical conformation, (b) the charge or hydrophobicity of the molecule at the site of the substitution, and/or (c) the bulk of the side chain.

A "deletion" refers to a change in the amino acid or nucleotide sequence that results in the

absence of one or more amino acid residues or nucleotides.

The term "derivative" refers to a chemically modified polynucleotide or polypeptide. Chemical modifications of a polynucleotide sequence can include, for example, replacement of hydrogen by an alkyl, acyl, hydroxyl, or amino group. A derivative polynucleotide encodes a polypeptide which retains at least one biological or immunological function of the natural molecule. A derivative polypeptide is one modified by glycosylation, pegylation, or any similar process that retains at least one biological or immunological function of the polypeptide from which it was derived.

A "detectable label" refers to a reporter molecule or enzyme that is capable of generating a measurable signal and is covalently or noncovalently joined to a polynucleotide or polypeptide.

A "fragment" is a unique portion of FLEXHT or the polynucleotide encoding FLEXHT which is identical in sequence to but shorter in length than the parent sequence. A fragment may comprise up to the entire length of the defined sequence, minus one nucleotide/amino acid residue. For example, a fragment may comprise from 5 to 1000 contiguous nucleotides or amino acid residues. A fragment used as a probe, primer, antigen, therapeutic molecule, or for other purposes, may be at least 5, 10, 15, 16, 20, 25, 30, 40, 50, 60, 75, 100, 150, 250 or at least 500 contiguous nucleotides or amino acid residues in length. Fragments may be preferentially selected from certain regions of a molecule. For example, a polypeptide fragment may comprise a certain length of contiguous amino acids selected from the first 250 or 500 amino acids (or first 25% or 50% of a polypeptide) as shown in a certain defined sequence. Clearly these lengths are exemplary, and any length that is supported by the specification, including the Sequence Listing, tables, and figures, may be encompassed by the present embodiments.

A fragment of SEQ ID NO:56-110 comprises a region of unique polynucleotide sequence that specifically identifies SEQ ID NO:56-110, for example, as distinct from any other sequence in the genome from which the fragment was obtained. A fragment of SEQ ID NO:56-110 is useful, for example, in hybridization and amplification technologies and in analogous methods that distinguish SEQ ID NO:56-110 from related polynucleotide sequences. The precise length of a fragment of SEQ ID NO:56-110 and the region of SEQ ID NO:56-110 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

A fragment of SEQ ID NO:1-55 is encoded by a fragment of SEQ ID NO:56-110. A fragment of SEQ ID NO:1-55 comprises a region of unique amino acid sequence that specifically identifies SEQ ID NO:1-55. For example, a fragment of SEQ ID NO:1-55 is useful as an immunogenic peptide for the development of antibodies that specifically recognize SEQ ID NO:1-55. The precise length of a fragment of SEQ ID NO:1-55 and the region of SEQ ID NO:1-55 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

A "full-length" polynucleotide sequence is one containing at least a translation initiation codon (e.g., methionine) followed by an open reading frame and a translation termination codon. A "full-length" polynucleotide sequence encodes a "full-length" polypeptide sequence.

"Homology" refers to sequence similarity or, interchangeably, sequence identity, between two or more polynucleotide sequences or two or more polypeptide sequences.

The terms "percent identity" and "% identity," as applied to polynucleotide sequences, refer to the percentage of residue matches between at least two polynucleotide sequences aligned using a standardized algorithm. Such an algorithm may insert, in a standardized and reproducible way, gaps in the sequences being compared in order to optimize alignment between two sequences, and therefore achieve a more meaningful comparison of the two sequences.

Percent identity between polynucleotide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program. This program is part of the LASERGENE software package, a suite of molecular biological analysis programs (DNASTAR, Madison WI). CLUSTAL V is described in Higgins, D.G. and P.M. Sharp (1989) CABIOS 5:151-153 and in Higgins, D.G. et al. (1992) CABIOS 8:189-191. For pairwise alignments of polynucleotide sequences, the default parameters are set as follows: Ktuple=2, gap penalty=5, window=4, and "diagonals saved"=4. The "weighted" residue weight table is selected as the default. Percent identity is reported by CLUSTAL V as the "percent similarity" between aligned polynucleotide sequences.

Alternatively, a suite of commonly used and freely available sequence comparison algorithms is provided by the National Center for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (BLAST) (Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410), which is available from several sources, including the NCBI, Bethesda, MD, and on the Internet at http://www.ncbi.nlm.nih.gov/BLAST/. The BLAST software suite includes various sequence analysis programs including "blastn," that is used to align a known polynucleotide sequence with other polynucleotide sequences from a variety of databases. Also available is a tool called "BLAST 2 Sequences" that is used for direct pairwise comparison of two nucleotide sequences. "BLAST 2 Sequences" can be accessed and used interactively at http://www.ncbi.nlm.nih.gov/gorf/bl2.html. The "BLAST 2 Sequences" tool can be used for both blastn and blastp (discussed below). BLAST programs are commonly used with gap and other parameters set to default settings. For example, to compare two nucleotide sequences, one may use blastn with the "BLAST 2 Sequences" tool Version 2.0.12 (April-21-2000) set at default parameters. Such default parameters may be, for example:

Matrix: BLOSUM62

Reward for match: 1

Penalty for mismatch: -2

Open Gap: 5 and Extension Gap: 2 penalties

Gap x drop-off: 50

Expect: 10

Word Size: 11

Filter: on

Percent identity may be measured over the length of an entire defined sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined sequence, for instance, a fragment of at least 20, at least 30, at least 40, at least 50, at least 70, at least 100, or at least 200 contiguous nucleotides. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures, or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

Nucleic acid sequences that do not show a high degree of identity may nevertheless encode similar amino acid sequences due to the degeneracy of the genetic code. It is understood that changes in a nucleic acid sequence can be made using this degeneracy to produce multiple nucleic acid sequences that all encode substantially the same protein.

The phrases "percent identity" and "% identity," as applied to polypeptide sequences, refer to the percentage of residue matches between at least two polypeptide sequences aligned using a standardized algorithm. Methods of polypeptide sequence alignment are well-known. Some alignment methods take into account conservative amino acid substitutions. Such conservative substitutions, explained in more detail above, generally preserve the charge and hydrophobicity at the site of substitution, thus preserving the structure (and therefore function) of the polypeptide.

Percent identity between polypeptide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program (described and referenced above). For pairwise alignments of polypeptide sequences using CLUSTAL V, the default parameters are set as follows: Ktuple=1, gap penalty=3, window=5, and "diagonals saved"=5. The PAM250 matrix is selected as the default residue weight table. As with polynucleotide alignments, the percent identity is reported by CLUSTAL V as the "percent similarity" between aligned polypeptide sequence pairs.

Alternatively the NCBI BLAST software suite may be used. For example, for a pairwise comparison of two polypeptide sequences, one may use the "BLAST 2 Sequences" tool Version 2.0.12 (Apr-21-2000) with blastp set at default parameters. Such default parameters may be, for example:

Matrix: BLOSUM62

Open Gap: 11 and Extension Gap: 1 penalties

Gap x drop-off: 50

Expect: 10

Word Size: 3

Filter: on

Percent identity may be measured over the length of an entire defined polypeptide sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined polypeptide sequence, for instance, a fragment of at least 15, at least 20, at least 30, at least 40, at least 50, at least 70 or at least 150 contiguous residues. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

"Human artificial chromosomes" (HACs) are linear microchromosomes which may contain DNA sequences of about 6 kb to 10 Mb in size, and which contain all of the elements required for chromosome replication, segregation and maintenance.

The term "humanized antibody" refers to an antibody molecule in which the amino acid sequence in the non-antigen binding regions has been altered so that the antibody more closely resembles a human antibody, and still retains its original binding ability.

"Hybridization" refers to the process by which a polynucleotide strand anneals with a complementary strand through base pairing under defined hybridization conditions. Specific hybridization is an indication that two nucleic acid sequences share a high degree of complementarity. Specific hybridization complexes form under permissive annealing conditions and remain hybridized after the "washing" step(s). The washing step(s) is particularly important in determining the stringency of the hybridization process, with more stringent conditions allowing less non-specific binding, i.e., binding between pairs of nucleic acid strands that are not perfectly matched. Permissive conditions for annealing of nucleic acid sequences are routinely determinable by one of ordinary skill in the art and may be consistent among hybridization experiments, whereas wash conditions may be varied among experiments to achieve the desired stringency, and therefore hybridization specificity. Permissive annealing conditions occur, for example, at 68°C in the presence of about 6 x SSC, about 1% (w/v) SDS, and about 100 µg/ml sheared, denatured salmon sperm DNA.

Generally, stringency of hybridization is expressed, in part, with reference to the temperature under which the wash step is carried out. Such wash temperatures are typically selected to be about 5°C to 20°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength and pH) at which 50% of the

target sequence hybridizes to a perfectly matched probe. An equation for calculating T_m and conditions for nucleic acid hybridization are well known and can be found in Sambrook, J. et al., 1989, Molecular Cloning: A Laboratory Manual, 2nd ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; specifically see volume 2, chapter 9.

High stringency conditions for hybridization between polynucleotides of the present invention include wash conditions of 68°C in the presence of about 0.2 x SSC and about 0.1% SDS, for 1 hour. Alternatively, temperatures of about 65°C, 60°C, 55°C, or 42°C may be used. SSC concentration may be varied from about 0.1 to 2 x SSC, with SDS being present at about 0.1%. Typically, blocking reagents are used to block non-specific hybridization. Such blocking reagents include, for instance, sheared and denatured salmon sperm DNA at about 100-200 μg/ml. Organic solvent, such as formamide at a concentration of about 35-50% v/v, may also be used under particular circumstances, such as for RNA:DNA hybridizations. Useful variations on these wash conditions will be readily apparent to those of ordinary skill in the art. Hybridization, particularly under high stringency conditions, may be suggestive of evolutionary similarity between the nucleotides. Such similarity is strongly indicative of a similar role for the nucleotides and their encoded polypeptides.

The term "hybridization complex" refers to a complex formed between two nucleic acid sequences by virtue of the formation of hydrogen bonds between complementary bases. A hybridization complex may be formed in solution (e.g., C_0t or R_0t analysis) or formed between one nucleic acid sequence present in solution and another nucleic acid sequence immobilized on a solid support (e.g., paper, membranes, filters, chips, pins or glass slides, or any other appropriate substrate to which cells or their nucleic acids have been fixed).

The words "insertion" and "addition" refer to changes in an amino acid or nucleotide sequence resulting in the addition of one or more amino acid residues or nucleotides, respectively.

"Immune response" can refer to conditions associated with inflammation, trauma, immune disorders, or infectious or genetic disease, etc. These conditions can be characterized by expression of various factors, e.g., cytokines, chemokines, and other signaling molecules, which may affect cellular and systemic defense systems.

An "immunogenic fragment" is a polypeptide or oligopeptide fragment of FLEXHT which is capable of eliciting an immune response when introduced into a living organism, for example, a mammal. The term "immunogenic fragment" also includes any polypeptide or oligopeptide fragment of FLEXHT which is useful in any of the antibody production methods disclosed herein or known in the art.

The term "microarray" refers to an arrangement of a plurality of polynucleotides, polypeptides, or other chemical compounds on a substrate.

The terms "element" and "array element" refer to a polynucleotide, polypeptide, or other chemical compound having a unique and defined position on a microarray.

The term "modulate" refers to a change in the activity of FLEXHT. For example, modulation may cause an increase or a decrease in protein activity, binding characteristics, or any other biological, functional, or immunological properties of FLEXHT.

The phrases "nucleic acid" and "nucleic acid sequence" refer to a nucleotide, oligonucleotide, polynucleotide, or any fragment thereof. These phrases also refer to DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA), or to any DNA-like or RNA-like material.

"Operably linked" refers to the situation in which a first nucleic acid sequence is placed in a functional relationship with a second nucleic acid sequence. For instance, a promoter is operably linked to a coding sequence if the promoter affects the transcription or expression of the coding sequence. Operably linked DNA sequences may be in close proximity or contiguous and, where necessary to join two protein coding regions, in the same reading frame.

"Peptide nucleic acid" (PNA) refers to an antisense molecule or anti-gene agent which comprises an oligonucleotide of at least about 5 nucleotides in length linked to a peptide backbone of amino acid residues ending in lysine. The terminal lysine confers solubility to the composition. PNAs preferentially bind complementary single stranded DNA or RNA and stop transcript elongation, and may be pegylated to extend their lifespan in the cell.

"Post-translational modification" of an FLEXHT may involve lipidation, glycosylation, phosphorylation, acetylation, racemization, proteolytic cleavage, and other modifications known in the art. These processes may occur synthetically or biochemically. Biochemical modifications will vary by cell type depending on the enzymatic milieu of FLEXHT.

"Probe" refers to nucleic acid sequences encoding FLEXHT, their complements, or fragments thereof, which are used to detect identical, allelic or related nucleic acid sequences. Probes are isolated oligonucleotides or polynucleotides attached to a detectable label or reporter molecule. Typical labels include radioactive isotopes, ligands, chemiluminescent agents, and enzymes. "Primers" are short nucleic acids, usually DNA oligonucleotides, which may be annealed to a target polynucleotide by complementary base-pairing. The primer may then be extended along the target DNA strand by a DNA polymerase enzyme. Primer pairs can be used for amplification (and identification) of a nucleic acid sequence, e.g., by the polymerase chain reaction (PCR).

Probes and primers as used in the present invention typically comprise at least 15 contiguous nucleotides of a known sequence. In order to enhance specificity, longer probes and primers may also be employed, such as probes and primers that comprise at least 20, 25, 30, 40, 50, 60, 70, 80, 90, 100, or at least 150 consecutive nucleotides of the disclosed nucleic acid sequences. Probes and primers may

be considerably longer than these examples, and it is understood that any length supported by the specification, including the tables, figures, and Sequence Listing, may be used.

Methods for preparing and using probes and primers are described in the references, for example Sambrook, J. et al., 1989, Molecular Cloning: A Laboratory Manual, 2nd ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; Ausubel, F.M. et al.,1987, Current Protocols in Molecular Biology, Greene Publ. Assoc. & Wiley-Intersciences, New York NY; Innis, M. et al., 1990, PCR Protocols, A Guide to Methods and Applications, Academic Press, San Diego CA. PCR primer pairs can be derived from a known sequence, for example, by using computer programs intended for that purpose such as Primer (Version 0.5, 1991, Whitehead Institute for Biomedical Research, Cambridge MA).

Oligonucleotides for use as primers are selected using software known in the art for such purpose. For example, OLIGO 4.06 software is useful for the selection of PCR primer pairs of up to 100 nucleotides each, and for the analysis of oligonucleotides and larger polynucleotides of up to 5,000 nucleotides from an input polynucleotide sequence of up to 32 kilobases. Similar primer selection programs have incorporated additional features for expanded capabilities. For example, the PrimOU primer selection program (available to the public from the Genome Center at University of Texas South West Medical Center, Dallas TX) is capable of choosing specific primers from megabase sequences and is thus useful for designing primers on a genome-wide scope. The Primer3 primer selection program (available to the public from the Whitehead Institute/MIT Center for Genome Research, Cambridge MA) allows the user to input a "mispriming library," in which sequences to avoid as primer binding sites are user-specified. Primer3 is useful, in particular, for the selection of oligonucleotides for microarrays. (The source code for the latter two primer selection programs may also be obtained from their respective sources and modified to meet the user's specific needs.) The PrimeGen program (available to the public from the UK Human Genome Mapping Project Resource Centre, Cambridge UK) designs primers based on multiple sequence alignments, thereby allowing selection of primers that hybridize to either the most conserved or least conserved regions of aligned nucleic acid sequences. Hence, this program is useful for identification of both unique and conserved oligonucleotides and polynucleotide fragments. The oligonucleotides and polynucleotide fragments identified by any of the above selection methods are useful in hybridization technologies, for example, as PCR or sequencing primers, microarray elements, or specific probes to identify fully or partially complementary polynucleotides in a sample of nucleic acids. Methods of oligonucleotide selection are not limited to those described above.

A "recombinant nucleic acid" is a sequence that is not naturally occurring or has a sequence that is made by an artificial combination of two or more otherwise separated segments of sequence.

This artificial combination is often accomplished by chemical synthesis or, more commonly, by the artificial manipulation of isolated segments of nucleic acids, e.g., by genetic engineering techniques such as those described in Sambrook, supra. The term recombinant includes nucleic acids that have been altered solely by addition, substitution, or deletion of a portion of the nucleic acid. Frequently, a recombinant nucleic acid may include a nucleic acid sequence operably linked to a promoter sequence. Such a recombinant nucleic acid may be part of a vector that is used, for example, to transform a cell.

Alternatively, such recombinant nucleic acids may be part of a viral vector, e.g., based on a vaccinia virus, that could be use to vaccinate a mammal wherein the recombinant nucleic acid is expressed, inducing a protective immunological response in the mammal.

A "regulatory element" refers to a nucleic acid sequence usually derived from untranslated regions of a gene and includes enhancers, promoters, introns, and 5' and 3' untranslated regions (UTRs). Regulatory elements interact with host or viral proteins which control transcription, translation, or RNA stability.

"Reporter molecules" are chemical or biochemical moieties used for labeling a nucleic acid, amino acid, or antibody. Reporter molecules include radionuclides; enzymes; fluorescent, chemiluminescent, or chromogenic agents; substrates; cofactors; inhibitors; magnetic particles; and other moieties known in the art.

An "RNA equivalent," in reference to a DNA sequence, is composed of the same linear sequence of nucleotides as the reference DNA sequence with the exception that all occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

The term "sample" is used in its broadest sense. A sample suspected of containing nucleic acids encoding FLEXHT, or fragments thereof, or FLEXHT itself, may comprise a bodily fluid; an extract from a cell, chromosome, organelle, or membrane isolated from a cell; a cell; genomic DNA, RNA, or cDNA, in solution or bound to a substrate; a tissue; a tissue print; etc.

The terms "specific binding" and "specifically binding" refer to that interaction between a protein or peptide and an agonist, an antibody, an antagonist, a small molecule, or any natural or synthetic binding composition. The interaction is dependent upon the presence of a particular structure of the protein, e.g., the antigenic determinant or epitope, recognized by the binding molecule. For example, if an antibody is specific for epitope "A," the presence of a polypeptide comprising the epitope A, or the presence of free unlabeled A, in a reaction containing free labeled A and the antibody will reduce the amount of labeled A that binds to the antibody.

The term "substantially purified" refers to nucleic acid or amino acid sequences that are removed from their natural environment and are isolated or separated, and are at least 60% free,

preferably at least 75% free, and most preferably at least 90% free from other components with which they are naturally associated.

A "substitution" refers to the replacement of one or more amino acid residues or nucleotides by different amino acid residues or nucleotides, respectively.

"Substrate" refers to any suitable rigid or semi-rigid support including membranes, filters, chips, slides, wafers, fibers, magnetic or nonmagnetic beads, gels, tubing, plates, polymers, microparticles and capillaries. The substrate can have a variety of surface forms, such as wells, trenches, pins, channels and pores, to which polynucleotides or polypeptides are bound.

A "transcript image" refers to the collective pattern of gene expression by a particular cell type or tissue under given conditions at a given time.

"Transformation" describes a process by which exogenous DNA is introduced into a recipient cell. Transformation may occur under natural or artificial conditions according to various methods well known in the art, and may rely on any known method for the insertion of foreign nucleic acid sequences into a prokaryotic or eukaryotic host cell. The method for transformation is selected based on the type of host cell being transformed and may include, but is not limited to, bacteriophage or viral infection, electroporation, heat shock, lipofection, and particle bombardment. The term "transformed" cells includes stably transformed cells in which the inserted DNA is capable of replication either as an autonomously replicating plasmid or as part of the host chromosome, as well as transiently transformed cells which express the inserted DNA or RNA for limited periods of time.

A "transgenic organism," as used herein, is any organism, including but not limited to animals and plants, in which one or more of the cells of the organism contains heterologous nucleic acid introduced by way of human intervention, such as by transgenic techniques well known in the art. The nucleic acid is introduced into the cell, directly or indirectly by introduction into a precursor of the cell, by way of deliberate genetic manipulation, such as by microinjection or by infection with a recombinant virus. The term genetic manipulation does not include classical cross-breeding, or in vitro fertilization, but rather is directed to the introduction of a recombinant DNA molecule. The transgenic organisms contemplated in accordance with the present invention include bacteria, cyanobacteria, fungi, plants, and animals. The isolated DNA of the present invention can be introduced into the host by methods known in the art, for example infection, transfection, transformation or transconjugation. Techniques for transferring the DNA of the present invention into such organisms are widely known and provided in references such as Sambrook et al. (1989), supra.

A "variant" of a particular nucleic acid sequence is defined as a nucleic acid sequence having at least 40% sequence identity to the particular nucleic acid sequence over a certain length of one of the nucleic acid sequences using blastn with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999)

set at default parameters. Such a pair of nucleic acids may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 95% or at least 98% or greater sequence identity over a certain defined length. A variant may be described as, for example, an "allelic" (as defined above), "splice," "species," or "polymorphic" variant. A splice variant may have significant identity to a reference molecule, but will generally have a greater or lesser number of polynucleotides due to alternative splicing of exons during mRNA processing. The corresponding polypeptide may possess additional functional domains or lack domains that are present in the reference molecule. Species variants are polynucleotide sequences that vary from one species to another. The resulting polypeptides generally will have significant amino acid identity relative to each other. A polymorphic variant is a variation in the polynucleotide sequence of a particular gene between individuals of a given species. Polymorphic variants also may encompass "single nucleotide polymorphisms" (SNPs) in which the polynucleotide sequence varies by one nucleotide base. The presence of SNPs may be indicative of, for example, a certain population, a disease state, or a propensity for a disease state.

A "variant" of a particular polypeptide sequence is defined as a polypeptide sequence having at least 40% sequence identity to the particular polypeptide sequence over a certain length of one of the polypeptide sequences using blastp with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of polypeptides may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, or at least 98% or greater sequence identity over a certain defined length of one of the polypeptides.

THE INVENTION

The invention is based on the discovery of new full-length molecules expressed in human tissues (FLEXHT), the polynucleotides encoding FLEXHT, and the use of these compositions for the analysis of human gene expression, genetic linkage, and genetic variability, and for the diagnosis, treatment, or prevention of developmental, cell proliferative, and immunological disorders.

Table 1 lists the Incyte clones used to assemble full length nucleotide sequences encoding FLEXHT. Columns 1 and 2 show the sequence identification numbers (SEQ ID NOs) of the polypeptide and nucleotide sequences, respectively. Column 3 shows the clone IDs of the Incyte clones in which nucleic acids encoding each FLEXHT were identified, and column 4 shows the cDNA libraries from which these clones were isolated. Column 5 shows Incyte clones and their corresponding cDNA libraries. Clones for which cDNA libraries are not indicated were derived from pooled cDNA libraries. The Incyte clones in column 5 were used to assemble the consensus nucleotide sequence of each FLEXHT and are useful as fragments in hybridization technologies.

The columns of Table 2 show various properties of each of the polypeptides of the invention: column 1 references the SEQ ID NO; column 2 shows the number of amino acid residues in each polypeptide; column 3 shows potential phosphorylation sites; and column 4 shows potential

glycosylation sites.

The columns of Table 3 show the tissue-specificity and diseases, disorders, or conditions associated with nucleotide sequences encoding FLEXHT. The first column of Table 3 lists the nucleotide SEQ ID NOs. Column 2 lists fragments of the nucleotide sequences of column 1. These fragments are useful, for example, in hybridization or amplification technologies to identify SEQ ID NO:56-110 and to distinguish between SEQ ID NO:56-110 and related polynucleotide sequences. The polypeptides encoded by the indicated fragments of SEQ ID NO:56, SEQ ID NO:58-69, SEQ ID NO:71-72, SEQ ID NO:74-76, SEQ ID NO:79-85, SEQ ID NO:87, SEQ ID NO:90-98, SEQ ID NO:100-103, and SEQ ID NO:105-110 are useful, for example, as immunogenic peptides. Column 3 lists tissue categories which express FLEXHT as a fraction of total tissues expressing FLEXHT. Column 4 lists diseases, disorders, or conditions associated with those tissues expressing FLEXHT as a fraction of total tissues expressing FLEXHT. Column 5 lists the vectors used to subclone each cDNA library.

The columns of Table 4 show descriptions of the tissues used to construct the cDNA libraries from which cDNA clones encoding FLEXHT were isolated. Column 1 references the nucleotide SEQ ID NOs, column 2 shows the cDNA libraries from which these clones were isolated, and column 3 shows the tissue origins and other descriptive information relevant to the cDNA libraries in column 2.

The columns of Table 5 show various properties of polypeptides of the invention: column 1 references the SEQ ID NO; column 2 shows the identity of the polypeptide along with relevant citations, all of which are expressly incorporated by reference herein in their entirety; column 3 shows the amino acid residues comprising signature sequences, domains, and motifs, and homologous sequences as identified by BLAST analysis; and column 4 shows analytical methods and in some cases, searchable databases to which the analytical methods were applied. The methods of column 4 were used to characterize each polypeptide through sequence homology and protein motifs.

The invention incorporates the nucleic acid sequences disclosed in the Sequence Listing and the use of these sequences in the diagnosis and treatment of disease states characterized by altered expression of FLEXHT genes or defects in FLEXHT protein function. The invention further utilizes these sequences in hybridization and amplification technologies, and in particular, in technologies which comprehensively assess gene expression patterns correlated with specific cells or tissues and their responses in vivo or in vitro to pharmaceutical agents, toxins, and other treatments. In this manner, the sequences of the present invention are used to develop a transcript image for a particular cell or tissue.

The invention also encompasses FLEXHT variants. A preferred FLEXHT variant is one which has at least about 80%, or alternatively at least about 90%, or even at least about 95% amino acid sequence identity to the FLEXHT amino acid sequence, and which contains at least one functional or

structural characteristic of FLEXHT.

The invention also encompasses polynucleotides which encode FLEXHT. In a particular embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:56-110, which encodes FLEXHT. The polynucleotide sequences of SEQ ID NO:56-110, as presented in the Sequence Listing, embrace the equivalent RNA sequences, wherein occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

The invention also encompasses a variant of a polynucleotide sequence encoding FLEXHT. In particular, such a variant polynucleotide sequence will have at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to the polynucleotide sequence encoding FLEXHT. A particular aspect of the invention encompasses a variant of a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:56-110 which has at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to a nucleic acid sequence selected from the group consisting of SEQ ID NO:56-110. Any one of the polynucleotide variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of FLEXHT.

It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of polynucleotide sequences encoding FLEXHT, some bearing minimal similarity to the polynucleotide sequences of any known and naturally occurring gene, may be produced. Thus, the invention contemplates each and every possible variation of polynucleotide sequence that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the polynucleotide sequence of naturally occurring FLEXHT, and all such variations are to be considered as being specifically disclosed.

Although nucleotide sequences which encode FLEXHT and its variants are generally capable of hybridizing to the nucleotide sequence of the naturally occurring FLEXHT under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding FLEXHT or its derivatives possessing a substantially different codon usage, e.g., inclusion of non-naturally occurring codons. Codons may be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding FLEXHT and its derivatives without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

The invention also encompasses production of DNA sequences which encode FLEXHT and

FLEXHT derivatives, or fragments thereof, entirely by synthetic chemistry. After production, the synthetic sequence may be inserted into any of the many available expression vectors and cell systems using reagents well known in the art. Moreover, synthetic chemistry may be used to introduce mutations into a sequence encoding FLEXHT or any fragment thereof.

Hybridization

Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing to the claimed polynucleotide sequences, and, in particular, to those shown in SEQ ID NO:56-110 and fragments thereof under various conditions of stringency. (See, e.g., Wahl, G.M. and S.L. Berger (1987) Methods Enzymol. 152:399-407; Kimmel, A.R. (1987) Methods Enzymol. 152:507-511.) For example, stringent salt concentration will ordinarily be less than about 750 mM NaCl and 75 mM trisodium citrate, preferably less than about 500 mM NaCl and 50 mM trisodium citrate, and most preferably less than about 250 mM NaCl and 25 mM trisodium citrate. Low stringency hybridization can be obtained in the absence of organic solvent, e.g., formamide, while high stringency hybridization can be obtained in the presence of at least about 35% formamide, and most preferably at least about 50% formamide. Stringent temperature conditions will ordinarily include temperatures of at least about 30°C, more preferably of at least about 37°C, and most preferably of at least about 42°C. Varying additional parameters, such as hybridization time, the concentration of detergent, e.g., sodium dodecyl sulfate (SDS), and the inclusion or exclusion of carrier DNA, are well known to those skilled in the art. Various levels of stringency are accomplished by combining these various conditions as needed. In a preferred embodiment, hybridization will occur at 30°C in 750 mM NaCl, 75 mM trisodium citrate, and 1% SDS. In a more preferred embodiment, hybridization will occur at 37°C in 500 mM NaCl, 50 mM trisodium citrate, 1% SDS, 35% formamide, and 100 μg/ml denatured salmon sperm DNA (ssDNA). In a most preferred embodiment, hybridization will occur at 42°C in 250 mM NaCl, 25 mM trisodium citrate, 1% SDS, 50 % formamide, and 200 μg/ml ssDNA. Useful variations on these conditions will be readily apparent to those skilled in the art.

The washing steps which follow hybridization can also vary in stringency. Wash stringency conditions can be defined by salt concentration and by temperature. As above, wash stringency can be increased by decreasing salt concentration or by increasing temperature. For example, stringent salt concentration for the wash steps will preferably be less than about 30 mM NaCl and 3 mM trisodium citrate, and most preferably less than about 15 mM NaCl and 1.5 mM trisodium citrate. Stringent temperature conditions for the wash steps will ordinarily include temperature of at least about 25°C, more preferably of at least about 42°C, and most preferably of at least about 68°C. In a preferred embodiment, wash steps will occur at 25°C in 30 mM NaCl, 3 mM trisodium citrate, and 0.1% SDS. In a more preferred embodiment, wash steps will occur at 42°C in 15 mM NaCl, 1.5 mM trisodium

citrate, and 0.1% SDS. In a most preferred embodiment, wash steps will occur at 68°C in 15 mM NaCl, 1.5 mM trisodium citrate, and 0.1% SDS. Additional variations on these conditions will be readily apparent to those skilled in the art. Further examples of hybridization conditions, including annealing and wash conditions, are described in "Definitions."

cDNA Sequencing

Methods for DNA sequencing are well known in the art and may be used to practice any of the embodiments of the invention. The methods may employ such enzymes as the Klenow fragment of DNA polymerase I, SEQUENASE (US Biochemical, Cleveland OH), Taq polymerase (Perkin-Elmer), thermostable T7 polymerase (Amersham Pharmacia Biotech, Piscataway NJ), or combinations of polymerases and proofreading exonucleases such as those found in the ELONGASE amplification system (Life Technologies, Gaithersburg MD). Preferably, sequence preparation is automated with machines such as the MICROLAB 2200 liquid transfer system (Hamilton, Reno NV), PTC200 thermal cycler (MJ Research, Watertown MA) and ABI CATALYST 800 thermal cycler (Perkin-Elmer). Sequencing is then carried out using either the ABI 373 or 377 DNA sequencing system (Perkin-Elmer), the MEGABACE 1000 DNA sequencing system (Molecular Dynamics, Sunnyvale CA), or other systems known in the art. The resulting sequences are analyzed using a variety of algorithms which are well known in the art. (See, e.g., Ausubel, F.M. (1997) Short Protocols in Molecular Biology, John Wiley & Sons, New York NY, unit 7.7; Meyers, R.A. (1995) Molecular Biology and Biotechnology, Wiley VCH, New York NY, pp. 856-853.)

The nucleic acid sequences encoding FLEXHT may be extended utilizing a partial nucleotide sequence and employing various PCR-based methods known in the art to detect upstream sequences, such as promoters and regulatory elements. For example, one method which may be employed, restriction-site PCR, uses universal and nested primers to amplify unknown sequence from genomic DNA within a cloning vector. (See, e.g., Sarkar, G. (1993) PCR Methods Applic. 2:318-322.) Another method, inverse PCR, uses primers that extend in divergent directions to amplify unknown sequence from a circularized template. The template is derived from restriction fragments comprising a known genomic locus and surrounding sequences. (See, e.g., Triglia, T. et al. (1988) Nucleic Acids Res. 16:8186.) A third method, capture PCR, involves PCR amplification of DNA fragments adjacent to known sequences in human and yeast artificial chromosome DNA. (See, e.g., Lagerstrom, M. et al. (1991) PCR Methods Applic. 1:111-119.) In this method, multiple restriction enzyme digestions and ligations may be used to insert an engineered double-stranded sequence into a region of unknown sequence before performing PCR. Other methods which may be used to retrieve unknown sequences are known in the art. (See, e.g., Parker, J.D. et al. (1991) Nucleic Acids Res. 19:3055-3060). Additionally, one may use PCR, nested primers, and PROMOTERFINDER libraries (Clontech, Palo

Alto CA) to walk genomic DNA. This procedure avoids the need to screen libraries and is useful in finding intron/exon junctions. For all PCR-based methods, primers may be designed using commercially available software, such as OLIGO 4.06 Primer Analysis software (National Biosciences, Plymouth MN) or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the template at temperatures of about 68°C to 72°C.

When screening for full-length cDNAs, it is preferable to use libraries that have been size-selected to include larger cDNAs. In addition, random-primed libraries, which often include sequences containing the 5' regions of genes, are preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries may be useful for extension of sequence into 5' non-transcribed regulatory regions.

Capillary electrophoresis systems which are commercially available may be used to analyze the size or confirm the nucleotide sequence of sequencing or PCR products. In particular, capillary sequencing may employ flowable polymers for electrophoretic separation, four different nucleotide-specific, laser-stimulated fluorescent dyes, and a charge coupled device camera for detection of the emitted wavelengths. Output/light intensity may be converted to electrical signal using appropriate software (e.g., GENOTYPER and SEQUENCE NAVIGATOR, Perkin-Elmer), and the entire process from loading of samples to computer analysis and electronic data display may be computer controlled. Capillary electrophoresis is especially preferable for sequencing small DNA fragments which may be present in limited amounts in a particular sample.

cDNA Expression

In another embodiment of the invention, polynucleotide sequences or fragments thereof which encode FLEXHT may be cloned in recombinant DNA molecules that direct expression of FLEXHT, or fragments or functional equivalents thereof, in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be produced and used to express FLEXHT.

The nucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter FLEXHT-encoding sequences for a variety of purposes including, but not limited to, modification of the cloning, processing, and/or expression of the gene product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. For example, oligonucleotide-mediated site-directed mutagenesis may be used to introduce mutations that create new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, and so forth.

The nucleotides of the present invention may be subjected to DNA shuffling techniques such

as MOLECULARBREEDING (Maxygen Inc., Santa Clara CA; described in U.S. Patent Number 5,837,458; Chang, C.-C. et al. (1999) Nat. Biotechnol. 17:793-797; Christians, F.C. et al. (1999) Nat. Biotechnol. 17:259-264; and Crameri, A. et al. (1996) Nat. Biotechnol. 14:315-319) to alter or improve the biological properties of FLEXHT, such as its biological or enzymatic activity or its ability to bind to other molecules or compounds. DNA shuffling is a process by which a library of gene variants is produced using PCR-mediated recombination of gene fragments. The library is then subjected to selection or screening procedures that identify those gene variants with the desired properties. These preferred variants may then be pooled and further subjected to recursive rounds of DNA shuffling and selection/screening. Thus, genetic diversity is created through "artificial" breeding and rapid molecular evolution. For example, fragments of a single gene containing random point mutations may be recombined, screened, and then reshuffled until the desired properties are optimized. Alternatively, fragments of a given gene may be recombined with fragments of homologous genes in the same gene family, either from the same or different species, thereby maximizing the genetic diversity of multiple naturally occurring genes in a directed and controllable manner.

In another embodiment, sequences encoding FLEXHT may be synthesized, in whole or in part, using chemical methods well known in the art. (See, e.g., Caruthers, M.H. et al. (1980) Nucleic Acids Symp. Ser. 7:215-223; and Horn, T. et al. (1980) Nucleic Acids Symp. Ser. 7:225-232.) Alternatively, FLEXHT itself or a fragment thereof may be synthesized using chemical methods. For example, peptide synthesis can be performed using various solution-phase or solid-phase techniques. (See, e.g., Creighton, T. (1984) Proteins, Structures and Molecular Properties, WH Freeman, New York NY, pp. 55-60; and Roberge, J.Y. et al. (1995) Science 269:202-204.) Automated synthesis may be achieved using the ABI 431A peptide synthesizer (Perkin-Elmer). Additionally, the amino acid sequence of FLEXHT, or any part thereof, may be altered during direct synthesis and/or combined with sequences from other proteins, or any part thereof, to produce a variant polypeptide or a polypeptide having a sequence of a naturally occurring polypeptide.

The peptide may be substantially purified by preparative high performance liquid chromatography. (See, e.g., Chiez, R.M. and F.Z. Regnier (1990) Methods Enzymol. 182:392-421.) The composition of the synthetic peptides may be confirmed by amino acid analysis or by sequencing. (See, e.g., Creighton, supra, pp. 28-53.)

In order to express a biologically active FLEXHT, the nucleotide sequences encoding FLEXHT or derivatives thereof may be inserted into an appropriate expression vector, i.e., a vector which contains the necessary elements for transcriptional and translational control of the inserted coding sequence in a suitable host. These elements include regulatory sequences, such as enhancers, constitutive and inducible promoters, and 5' and 3' untranslated regions in the vector and in

polynucleotide sequences encoding FLEXHT. Such elements may vary in their strength and specificity. Specific initiation signals may also be used to achieve more efficient translation of sequences encoding FLEXHT. Such signals include the ATG initiation codon and adjacent sequences, e.g. the Kozak sequence. In cases where sequences encoding FLEXHT and its initiation codon and upstream regulatory sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals including an in-frame ATG initiation codon should be provided by the vector. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers appropriate for the particular host cell system used. (See, e.g., Scharf, D. et al. (1994) Results Probl. Cell Differ. 20:125-162.)

Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding FLEXHT and appropriate transcriptional and translational control elements. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. (See, e.g., Sambrook, J. et al. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview NY, ch. 4, 8, and 16-17; Ausubel, F.M. et al. (1995) Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, ch. 9, 13, and 16.)

A variety of expression vector/host systems may be utilized to contain and express sequences encoding FLEXHT. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with viral expression vectors (e.g., baculovirus); plant cell systems transformed with viral expression vectors (e.g., cauliflower mosaic virus, CaMV, or tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal cell systems. (See, e.g., Sambrook, supra; Ausubel, supra; Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509; Bitter, G.A. et al. (1987) Methods Enzymol. 153:516-544; Scorer, C.A. et al. (1994) Bio/Technology 12:181-184; Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945; Takamatsu, N. (1987) EMBO J. 6:307-311; Coruzzi, G. et al. (1984) EMBO J. 3:1671-1680; Broglie, R. et al. (1984) Science 224:838-843; Winter, J. et al. (1991) Results Probl. Cell Differ. 17:85-105; The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196; Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659; and Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355.) Expression vectors derived from retroviruses, adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of nucleotide sequences to the targeted organ, tissue, or cell population. (See, e.g., Di

Nicola, M. et al. (1998) Cancer Gen. Ther. 5(6):350-356; Yu, M. et al., (1993) Proc. Natl. Acad. Sci. USA 90(13):6340-6344; Buller, R.M. et al. (1985) Nature 317(6040):813-815; McGregor, D.P. et al. (1994) Mol. Immunol. 31(3):219-226; and Verma, I.M. and N. Somia (1997) Nature 389:239-242.) The invention is not limited by the host cell employed.

In bacterial systems, a number of cloning and expression vectors may be selected depending upon the use intended for polynucleotide sequences encoding FLEXHT. For example, routine cloning, subcloning, and propagation of polynucleotide sequences encoding FLEXHT can be achieved using a multifunctional E. coli vector such as PBLUESCRIPT (Stratagene, La Jolla CA) or PSPORT1 plasmid (Life Technologies). Ligation of sequences encoding FLEXHT into the vector's multiple cloning site disrupts the *lacZ* gene, allowing a colorimetric screening procedure for identification of transformed bacteria containing recombinant molecules. In addition, these vectors may be useful for in vitro transcription, dideoxy sequencing, single strand rescue with helper phage, and creation of nested deletions in the cloned sequence. (See, e.g., Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509.) When large quantities of FLEXHT are needed, e.g. for the production of antibodies, vectors which direct high level expression of FLEXHT may be used. For example, vectors containing the strong, inducible T5 or T7 bacteriophage promoter may be used.

Yeast expression systems may be used for production of FLEXHT. A number of vectors containing constitutive or inducible promoters, such as alpha factor, alcohol oxidase, and PGH promoters, may be used in the yeast Saccharomyces cerevisiae or Pichia pastoris. In addition, such vectors direct either the secretion or intracellular retention of expressed proteins and enable integration of foreign sequences into the host genome for stable propagation. (See, e.g., Ausubel, 1995, supra; Bitter, supra; and Scorer, supra.)

Plant systems may also be used for expression of FLEXHT. Transcription of sequences encoding FLEXHT may be driven viral promoters, e.g., the 35S and 19S promoters of CaMV used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) EMBO J. 6:307-311). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used. (See, e.g., Coruzzi, supra; Broglie, supra; and Winter, supra.) These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. (See, e.g., The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196.)

In mammalian cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, sequences encoding FLEXHT may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain

infective virus which expresses FLEXHT in host cells. (See, e.g., Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659.) In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells. SV40 or EBV-based vectors may also be used for high-level protein expression.

Human artificial chromosomes (HACs) may also be employed to deliver larger fragments of DNA than can be contained in and expressed from a plasmid. HACs of about 6 kb to 10 Mb are constructed and delivered via conventional delivery methods (liposomes, polycationic amino polymers, or vesicles) for therapeutic purposes. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355.)

For long term production of recombinant proteins in mammalian systems, stable expression of FLEXHT in cell lines is preferred. For example, sequences encoding FLEXHT can be transformed into cell lines using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for about 1 to 2 days in enriched media before being switched to selective media. The purpose of the selectable marker is to confer resistance to a selective agent, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be propagated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase genes, for use in *tk⁻* and *apr⁻* cells, respectively. (See, e.g., Wigler, M. et al. (1977) Cell 11:223-232; Lowy, I. et al. (1980) Cell 22:817-823.) Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate; *neo* confers resistance to the aminoglycosides neomycin and G-418; and *als* and *pat* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively. (See, e.g., Wigler, M. et al. (1980) Proc. Natl. Acad. Sci. USA 77:3567-3570; Colbere-Garapin, F. et al. (1981) J. Mol. Biol. 150:1-14.) Additional selectable genes have been described, e.g., *trpB* and *hisD*, which alter cellular requirements for metabolites. (See, e.g., Hartman, S.C. and R.C. Mulligan (1988) Proc. Natl. Acad. Sci. USA 85:8047-8051.) Visible markers, e.g., anthocyanins, green fluorescent proteins (GFP; Clontech), ß glucuronidase and its substrate ß-glucuronide, or luciferase and its substrate luciferin may be used. These markers can be used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system. (See, e.g., Rhodes, C.A. (1995) Methods Mol. Biol. 55:121-131.)

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, the presence and expression of the gene may need to be confirmed. For example, if the

sequence encoding FLEXHT is inserted within a marker gene sequence, transformed cells containing sequences encoding FLEXHT can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding FLEXHT under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

In general, host cells that contain the nucleic acid sequence encoding FLEXHT and that express FLEXHT may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations, PCR amplification, and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein sequences.

Immunological methods for detecting and measuring the expression of FLEXHT using either specific polyclonal or monoclonal antibodies are known in the art. Examples of such techniques include enzyme-linked immunosorbent assays (ELISAs), radioimmunoassays (RIAs), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on FLEXHT is preferred, but a competitive binding assay may be employed. These and other assays are well known in the art. (See, e.g., Hampton, R. et al. (1990) Serological Methods, a Laboratory Manual, APS Press, St. Paul MN, Sect. IV; Coligan, J.E. et al. (1997) Current Protocols in Immunology, Greene Pub. Associates and Wiley-Interscience, New York NY; and Pound, J.D. (1998) Immunochemical Protocols, Humana Press, Totowa NJ.)

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding FLEXHT include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, the sequences encoding FLEXHT, or any fragments thereof, may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits, such as those provided by Amersham Pharmacia Biotech, Promega (Madison WI), and US Biochemical. Suitable reporter molecules or labels which may be used for ease of detection include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with nucleotide sequences encoding FLEXHT may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a transformed cell may be secreted or retained intracellularly depending on the sequence

and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode FLEXHT may be designed to contain signal sequences which direct secretion of FLEXHT through a prokaryotic or eukaryotic cell membrane.

In addition, a host cell strain may be chosen for its ability to modulate expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" or "pro" form of the protein may also be used to specify protein targeting, folding, and/or activity. Different host cells which have specific cellular machinery and characteristic mechanisms for post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and WI38) are available from the American Type Culture Collection (ATCC, Manassas VA) and may be chosen to ensure the correct modification and processing of the foreign protein.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences encoding FLEXHT may be ligated to a heterologous sequence resulting in translation of a fusion protein in any of the aforementioned host systems. For example, a chimeric FLEXHT protein containing a heterologous moiety that can be recognized by a commercially available antibody may facilitate the screening of peptide libraries for inhibitors of FLEXHT activity. Heterologous protein and peptide moieties may also facilitate purification of fusion proteins using commercially available affinity matrices. Such moieties include, but are not limited to, glutathione S-transferase (GST), maltose binding protein (MBP), thioredoxin (Trx), calmodulin binding peptide (CBP), 6-His, FLAG, *c-myc*, and hemagglutinin (HA). GST, MBP, Trx, CBP, and 6-His enable purification of their cognate fusion proteins on immobilized glutathione, maltose, phenylarsine oxide, calmodulin, and metal-chelate resins, respectively. FLAG, *c-myc*, and hemagglutinin (HA) enable immunoaffinity purification of fusion proteins using commercially available monoclonal and polyclonal antibodies that specifically recognize these epitope tags. A fusion protein may also be engineered to contain a proteolytic cleavage site located between the FLEXHT encoding sequence and the heterologous protein sequence, so that FLEXHT may be cleaved away from the heterologous moiety following purification. Methods for fusion protein expression and purification are discussed in Ausubel (1995, supra, ch. 10). A variety of commercially available kits may also be used to facilitate expression and purification of fusion proteins.

In a further embodiment of the invention, synthesis of radiolabeled FLEXHT may be achieved in vitro using the TNT rabbit reticulocyte lysate or wheat germ extract system (Promega). These systems couple transcription and translation of protein-coding sequences operably associated with the T7, T3, or SP6 promoters. Translation takes place in the presence of a radiolabeled amino acid precursor, for example, ^{35}S-methionine.

Fragments of FLEXHT may be produced not only by recombinant production, but also by direct peptide synthesis using solid-phase techniques. (See, e.g., Creighton, supra, pp. 55-60.) Protein synthesis may be performed by manual techniques or by automation. Automated synthesis may be achieved, for example, using the ABI 431A peptide synthesizer (Perkin-Elmer). Various fragments of FLEXHT may be synthesized separately and then combined to produce the full length molecule.

Screening Assays

FLEXHT of the present invention or fragments thereof may be used to screen for compounds that specifically bind to FLEXHT. At least one and up to a plurality of test compounds may be screened for specific binding to FLEXHT. Examples of test compounds include antibodies, oligonucleotides, proteins (e.g., receptors), or small molecules.

In one embodiment, the compound thus identified is closely related to the natural ligand of FLEXHT, e.g., a ligand or fragment thereof, a natural substrate, a structural or functional mimetic, or a natural binding partner. (See, Coligan, J.E. et al. (1991) Current Protocols in Immunology 1(2): Chapter 5.) Similarly, the compound can be closely related to the natural receptor to which FLEXHT binds, or to at least a fragment of the receptor, e.g., the ligand binding site. In either case, the compound can be rationally designed using known techniques. In one embodiment, screening for these compounds involves producing appropriate cells which express FLEXHT, either as a secreted protein or on the cell membrane. Preferred cells include cells from mammals, yeast, Drosophila, or E. coli. Cells expressing FLEXHT or cell membrane fractions which contain FLEXHT are then contacted with a test compound and binding, stimulation, or inhibition of activity of either FLEXHT or the compound is analyzed.

An assay may simply test binding of a test compound to the polypeptide, wherein binding is detected by a fluorophore, radioisotope, enzyme conjugate, or other detectable label. For example, the assay may comprise the steps of combining at least one test compound with FLEXHT, either in solution or affixed to a solid support, and detecting the binding of FLEXHT to the compound. Alternatively, the assay may detect or measure binding of a test compound in the presence of a labeled competitor. Additionally, the assay may be carried out using cell-free preparations, chemical libraries, or natural product mixtures, and the test compound(s) may be free in solution or affixed to a solid support.

FLEXHT of the present invention or fragments thereof may be used to screen for compounds that modulate the activity of FLEXHT. Such compounds may include agonists, antagonists, or partial or inverse agonists. In one embodiment, an assay is performed under conditions permissive for FLEXHT activity, wherein FLEXHT is combined with at least one test compound, and the activity of FLEXHT in the presence of a test compound is compared with the activity of FLEXHT in the absence of the test compound. A change in the activity of FLEXHT in the presence of the test compound is

indicative of a compound that modulates the activity of FLEXHT. Alternatively, a test compound is combined with an in vitro or cell-free system comprising FLEXHT under conditions suitable for FLEXHT activity, and the assay is performed. In either of these assays, a test compound which modulates the activity of FLEXHT may do so indirectly and need not come in direct contact with the test compound. At least one and up to a plurality of test compounds may be screened.

Preferably, an ELISA assay using, e.g., a monoclonal or polyclonal antibody, can measure polypeptide level in a sample. The antibody can measure polypeptide level by either binding, directly or indirectly, to the polypeptide or by competing with the polypeptide for a substrate.

All of the above assays can be used in a diagnostic or prognostic context. The molecules discovered using these assays can be used to treat disease or to bring about a particular result in a patient (e.g., blood vessel growth) by activating or inhibiting the polypeptide/molecule. Moreover, the assays can discover agents which may inhibit or enhance the production of the polypeptide from suitably manipulated cells or tissues.

Transcript Imaging

Another embodiment relates to the use of polynucleotide sequences encoding FLEXGEM to develop a transcript image of a tissue or cell type. A transcript image is the collective pattern of gene expression by a particular tissue or cell type under given conditions and at a given time. This pattern of gene expression is defined by the number of expressed genes and their abundance. Thus the polynucleotide sequences of the present invention may be used to develop a transcript image of a tissue or cell type by hybridizing, preferably in a microarray format, the polynucleotide sequences of the present invention to the totality of transcripts or reverse transcripts of a tissue or cell type. The resultant transcript image would provide a profile of FLEXGEM gene activity.

Transcript images which profile FLEXGEM gene expression may be generated using transcripts isolated from tissues, cell lines, biopsies, or other biological samples. The transcript image may thus reflect FLEXGEM gene expression in vivo, as in the case of a tissue or biopsy sample, or in vitro, as in the case of a cell line. Transcript images may be used to profile FLEXGEM gene expression in distinct tissue types. This process can be used to determine FLEXGEM gene activity in a particular tissue type relative to this activity in a different tissue type. Transcript images may be used to generate a profile of FLEXGEM gene expression characteristic of diseased tissue. Transcript images of tissues before and after treatment may be used for diagnostic purposes, to monitor the progression of disease, and to monitor the efficacy of drug treatments for diseases which affect the activity of genes encoding FLEXGEM.

Transcript images which profile FLEXGEM gene expression may also be used in conjunction with in vitro model systems and preclinical evaluation of pharmaceuticals. Transcript images of cell lines can be used to assess FLEXGEM activity and/or to identify cell lines that lack or misregulate

this activity. Such cell lines may then be treated with pharmaceutical agents, and a transcript image following treatment may indicate the efficacy of these agents in restoring desired levels of this activity. A similar approach may be used to assess the toxicity of pharmaceutical agents as reflected by undesirable changes in FLEXGEM activity. Candidate pharmaceutical agents may be evaluated by comparing their associated transcript images with those of pharmaceutical agents of known effectiveness.

Transgenic Animals

In another embodiment, polynucleotides encoding FLEXHT or their mammalian homologs may be "knocked out" in an animal model system using homologous recombination in embryonic stem (ES) cells. Such techniques are well known in the art and are useful for the generation of animal models of human disease. (See, e.g., U.S. Patent No. 5,175,383 and U.S. Patent No. 5,767,337.) For example, mouse ES cells, such as the mouse 129/SvJ cell line, are derived from the early mouse embryo and grown in culture. The ES cells are transformed with a vector containing the gene of interest disrupted by a marker gene, e.g., the neomycin phosphotransferase gene (neo; Capecchi, M.R. (1989) Science 244:1288-1292). The vector integrates into the corresponding region of the host genome by homologous recombination. Alternatively, homologous recombination takes place using the Cre-loxP system to knockout a gene of interest in a tissue- or developmental stage-specific manner (Marth, J.D. (1996) Clin. Invest. 97:1999-2002; Wagner, K.U. et al. (1997) Nucleic Acids Res. 25:4323-4330). Transformed ES cells are identified and microinjected into mouse cell blastocysts such as those from the C57BL/6 mouse strain. The blastocysts are surgically transferred to pseudopregnant dams, and the resulting chimeric progeny are genotyped and bred to produce heterozygous or homozygous strains. Transgenic animals thus generated may be tested with potential therapeutic or toxic agents.

Polynucleotides encoding FLEXHT may also be manipulated in vitro in ES cells derived from human blastocysts. Human ES cells have the potential to differentiate into at least eight separate cell lineages including endoderm, mesoderm, and ectodermal cell types. These cell lineages differentiate into, for example, neural cells, hematopoietic lineages, and cardiomyocytes (Thomson, J.A. et al. (1998) Science 282:1145-1147).

Polynucleotides encoding FLEXHT can also be used to create "knockin" humanized animals (pigs) or transgenic animals (mice or rats) to model human disease. With knockin technology, a region of a polynucleotide encoding FLEXHT is injected into animal ES cells, and the injected sequence integrates into the animal cell genome. Transformed cells are injected into blastulae, and the blastulae are implanted as described above. Transgenic progeny or inbred lines are studied and treated with potential pharmaceutical agents to obtain information on treatment of a human disease. Alternatively, a mammal inbred to overexpress FLEXHT, e.g., by secreting FLEXHT in its milk, may

also serve as a convenient source of that protein (Janne, J. et al. (1998) Biotechnol. Annu. Rev. 4:55-74).

THERAPEUTICS

Chemical and structural similarity, e.g., in the context of sequences and motifs, exists between regions of FLEXHT and full-length molecules expressed in human tissues. In addition, the expression of FLEXHT is closely associated with cell proliferation and the immune response. Therefore, FLEXHT appears to play a role in developmental, cell proliferative, and immunological disorders. In the treatment of disorders associated with increased FLEXHT expression or activity, it is desirable to decrease the expression or activity of FLEXHT. In the treatment of disorders associated with decreased FLEXHT expression or activity, it is desirable to increase the expression or activity of FLEXHT.

Therefore, in one embodiment, FLEXHT or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of FLEXHT. Examples of such disorders include, but are not limited to, a developmental disorder such as renal tubular acidosis, anemia, Cushing's syndrome, achondroplastic dwarfism, Duchenne and Becker muscular dystrophy, epilepsy, gonadal dysgenesis, WAGR syndrome (Wilms' tumor, aniridia, genitourinary abnormalities, and mental retardation), Smith-Magenis syndrome, myelodysplastic syndrome, hereditary mucoepithelial dysplasia, hereditary keratodermas, hereditary neuropathies such as Charcot-Marie-Tooth disease and neurofibromatosis, hypothyroidism, hydrocephalus, a seizure disorder such as Syndenham's chorea and cerebral palsy, spina bifida, anencephaly, craniorachischisis, congenital glaucoma, cataract, and sensorineural hearing loss; a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, a cancer of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; and an immunological disorder such as inflammation, actinic keratosis, acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, arteriosclerosis, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, bronchitis, bursitis, cholecystitis, cirrhosis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, paroxysmal nocturnal hemoglobinuria, hepatitis, hypereosinophilia, irritable bowel

syndrome, episodic lymphopenia with lymphocytotoxins, mixed connective tissue disease (MCTD), multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, myelofibrosis, osteoarthritis, osteoporosis, pancreatitis, polycythemia vera, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, primary thrombocythemia, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, trauma, and hematopoietic cancer including lymphoma, leukemia, and myeloma.

In another embodiment, a vector capable of expressing FLEXHT or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of FLEXHT including, but not limited to, those described above.

In a further embodiment, a pharmaceutical composition comprising a substantially purified FLEXHT in conjunction with a suitable pharmaceutical carrier may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of FLEXHT including, but not limited to, those provided above.

In still another embodiment, an agonist which modulates the activity of FLEXHT may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of FLEXHT including, but not limited to, those listed above.

In a further embodiment, an antagonist of FLEXHT may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of FLEXHT. Examples of such disorders include, but are not limited to, those developmental, cell proliferative, and immunological disorders described above. In one aspect, an antibody which specifically binds FLEXHT may be used directly as an antagonist or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissues which express FLEXHT.

In an additional embodiment, a vector expressing the complement of the polynucleotide encoding FLEXHT may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of FLEXHT including, but not limited to, those described above.

In other embodiments, any of the proteins, antagonists, antibodies, agonists, complementary sequences, or vectors of the invention may be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy may be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents may act synergistically to effect the treatment or prevention of the various disorders described above. Using this approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

An antagonist of FLEXHT may be produced using methods which are generally known in the

art. In particular, purified FLEXHT may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind FLEXHT. Antibodies to FLEXHT may also be generated using methods that are well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, and single chain antibodies, Fab fragments, and fragments produced by a Fab expression library. Neutralizing antibodies (i.e., those which inhibit dimer formation) are generally preferred for therapeutic use.

For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others may be immunized by injection with FLEXHT or with any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include, but are not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, KLH, and dinitrophenol. Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and Corynebacterium parvum are especially preferable.

It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies to FLEXHT have an amino acid sequence consisting of at least about 5 amino acids, and generally will consist of at least about 10 amino acids. It is also preferable that these oligopeptides, peptides, or fragments are identical to a portion of the amino acid sequence of the natural protein. Short stretches of FLEXHT amino acids may be fused with those of another protein, such as KLH, and antibodies to the chimeric molecule may be produced.

Monoclonal antibodies to FLEXHT may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique. (See, e.g., Kohler, G. et al. (1975) Nature 256:495-497; Kozbor, D. et al. (1985) J. Immunol. Methods 81:31-42; Cote, R.J. et al. (1983) Proc. Natl. Acad. Sci. USA 80:2026-2030; and Cole, S.P. et al. (1984) Mol. Cell Biol. 62:109-120.)

In addition, techniques developed for the production of "chimeric antibodies," such as the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used. (See, e.g., Morrison, S.L. et al. (1984) Proc. Natl. Acad. Sci. USA 81:6851-6855; Neuberger, M.S. et al. (1984) Nature 312:604-608; and Takeda, S. et al. (1985) Nature 314:452-454.) Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce FLEXHT-specific single chain antibodies. Antibodies with related specificity, but of distinct idiotypic composition, may be generated by chain shuffling from random combinatorial immunoglobulin libraries. (See, e.g., Burton, D.R. (1991) Proc. Natl. Acad. Sci. USA 88:10134-10137.)

Antibodies may also be produced by inducing in vivo production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature. (See, e.g., Orlandi, R. et al. (1989) Proc. Natl. Acad. Sci. USA 86:3833-3837; Winter, G. et al. (1991) Nature 349:293-299.)

Antibody fragments which contain specific binding sites for FLEXHT may also be generated. For example, such fragments include, but are not limited to, $F(ab')_2$ fragments produced by pepsin digestion of the antibody molecule and Fab fragments generated by reducing the disulfide bridges of the F(ab')2 fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity. (See, e.g., Huse, W.D. et al. (1989) Science 246:1275-1281.)

Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between FLEXHT and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering FLEXHT epitopes is generally used, but a competitive binding assay may also be employed (Pound, supra).

Various methods such as Scatchard analysis in conjunction with radioimmunoassay techniques may be used to assess the affinity of antibodies for FLEXHT. Affinity is expressed as an association constant, K_a, which is defined as the molar concentration of FLEXHT-antibody complex divided by the molar concentrations of free antigen and free antibody under equilibrium conditions. The K_a determined for a preparation of polyclonal antibodies, which are heterogeneous in their affinities for multiple FLEXHT epitopes, represents the average affinity, or avidity, of the antibodies for FLEXHT. The K_a determined for a preparation of monoclonal antibodies, which are monospecific for a particular FLEXHT epitope, represents a true measure of affinity. High-affinity antibody preparations with K_a ranging from about 10^9 to 10^{12} L/mole are preferred for use in immunoassays in which the FLEXHT-antibody complex must withstand rigorous manipulations. Low-affinity antibody preparations with K_a ranging from about 10^6 to 10^7 L/mole are preferred for use in immunopurification and similar procedures which ultimately require dissociation of FLEXHT, preferably in active form, from the antibody (Catty, D. (1988) Antibodies, Volume I: A Practical Approach, IRL Press, Washington DC; Liddell, J.E. and A. Cryer (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY).

The titer and avidity of polyclonal antibody preparations may be further evaluated to determine the quality and suitability of such preparations for certain downstream applications. For example, a

polyclonal antibody preparation containing at least 1-2 mg specific antibody/ml, preferably 5-10 mg specific antibody/ml, is generally employed in procedures requiring precipitation of FLEXHT-antibody complexes. Procedures for evaluating antibody specificity, titer, and avidity, and guidelines for antibody quality and usage in various applications, are generally available. (See, e.g., Catty, supra, and Coligan et al., supra.)

In another embodiment of the invention, the polynucleotides encoding FLEXHT, or any fragment or complement thereof, may be used for therapeutic purposes. In one aspect, modifications of gene expression can be achieved by designing complementary sequences or antisense molecules (DNA, RNA, PNA, or modified oligonucleotides) to the coding or regulatory regions of the gene encoding FLEXHT. Such technology is well known in the art, and antisense oligonucleotides or larger fragments can be designed from various locations along the coding or control regions of sequences encoding FLEXHT. (See, e.g., Agrawal, S., ed. (1996) Antisense Therapeutics, Humana Press Inc., Totawa NJ.)

In therapeutic use, any gene delivery system suitable for introduction of the antisense sequences into appropriate target cells can be used. Antisense sequences can be delivered intracellularly in the form of an expression plasmid which, upon transcription, produces a sequence complementary to at least a portion of the cellular sequence encoding the target protein. (See, e.g., Slater, J.E. et al. (1998) J. Allergy Clin. Immunol. 102(3):469-475; and Scanlon, K.J. et al. (1995) 9(13):1288-1296.) Antisense sequences can also be introduced intracellularly through the use of viral vectors, such as retrovirus and adeno-associated virus vectors. (See, e.g., Miller, A.D. (1990) Blood 76:271; Ausubel, supra; Uckert, W. and W. Walther (1994) Pharmacol. Ther. 63(3):323-347.) Other gene delivery mechanisms include liposome-derived systems, artificial viral envelopes, and other systems known in the art. (See, e.g., Rossi, J.J. (1995) Br. Med. Bull. 51(1):217-225; Boado, R.J. et al. (1998) J. Pharm. Sci. 87(11):1308-1315; and Morris, M.C. et al. (1997) Nucleic Acids Res. 25(14):2730-2736.)

In another embodiment of the invention, polynucleotides encoding FLEXHT may be used for somatic or germline gene therapy. Gene therapy may be performed to (i) correct a genetic deficiency (e.g., in the cases of severe combined immunodeficiency (SCID)-X1 disease characterized by X-linked inheritance (Cavazzana-Calvo, M. et al. (2000) Science 288:669-672), severe combined immunodeficiency syndrome associated with an inherited adenosine deaminase (ADA) deficiency (Blaese, R.M. et al. (1995) Science 270:475-480; Bordignon, C. et al. (1995) Science 270:470-475), cystic fibrosis (Zabner, J. et al. (1993) Cell 75:207-216; Crystal, R.G. et al. (1995) Hum. Gene Therapy 6:643-666; Crystal, R.G. et al. (1995) Hum. Gene Therapy 6:667-703), thalassamias, familial hypercholesterolemia, and hemophilia resulting from Factor VIII or Factor IX deficiencies (Crystal, R.G. (1995) Science 270:404-410; Verma, I.M. and Somia, N. (1997) Nature 389:239-242)), (ii)

express a conditionally lethal gene product (e.g., in the case of cancers which result from unregulated cell proliferation), or (iii) express a protein which affords protection against intracellular parasites (e.g., against human retroviruses, such as human immunodeficiency virus (HIV) (Baltimore, D. (1988) Nature 335:395-396; Poeschla, E. et al. (1996) Proc. Natl. Acad. Sci. USA. 93:11395-11399), hepatitis B or C virus (HBV, HCV); fungal parasites, such as Candida albicans and Paracoccidioides brasiliensis; and protozoan parasites such as Plasmodium falciparum and Trypanosoma cruzi). In the case where a genetic deficiency in FLEXHT expression or regulation causes disease, the expression of FLEXHT from an appropriate population of transduced cells may alleviate the clinical manifestations caused by the genetic deficiency.

In a further embodiment of the invention, diseases or disorders caused by deficiencies in FLEXHT are treated by constructing mammalian expression vectors encoding FLEXHT and introducing these vectors by mechanical means into FLEXHT-deficient cells. Mechanical transfer technologies for use with cells in vivo or ex vitro include (i) direct DNA microinjection into individual cells, (ii) ballistic gold particle delivery, (iii) liposome-mediated transfection, (iv) receptor-mediated gene transfer, and (v) the use of DNA transposons (Morgan, R.A. and W.F. Anderson (1993) Annu. Rev. Biochem. 62:191-217; Ivics, Z. (1997) Cell 91:501-510; Boulay, J-L. and H. Récipon (1998) Curr. Opin. Biotechnol. 9:445-450).

Expression vectors that may be effective for the expression of FLEXHT include, but are not limited to, the PCDNA 3.1, EPITAG, PRCCMV2, PREP, PVAX vectors (Invitrogen, Carlsbad CA), PCMV-SCRIPT, PCMV-TAG, PEGSH/PERV (Stratagene, La Jolla CA), and PTET-OFF, PTET-ON, PTRE2, PTRE2-LUC, PTK-HYG (Clontech, Palo Alto CA). FLEXHT may be expressed using (i) a constitutively active promoter, (e.g., from cytomegalovirus (CMV), Rous sarcoma virus (RSV), SV40 virus, thymidine kinase (TK), or β-actin genes), (ii) an inducible promoter (e.g., the tetracycline-regulated promoter (Gossen, M. and H. Bujard (1992) Proc. Natl. Acad. Sci. U.S.A. 89:5547-5551; Gossen, M. et al. (1995) Science 268:1766-1769; Rossi, F.M.V. and H.M. Blau (1998) Curr. Opin. Biotechnol. 9:451-456), commercially available in the T-REX plasmid (Invitrogen)); the ecdysone-inducible promoter (available in the plasmids PVGRXR and PIND; Invitrogen); the FK506/rapamycin inducible promoter; or the RU486/mifepristone inducible promoter (Rossi, F.M.V. and H.M. Blau, supra)), or (iii) a tissue-specific promoter or the native promoter of the endogenous gene encoding FLEXHT from a normal individual.

Commercially available liposome transformation kits (e.g., the PERFECT LIPID TRANSFECTION KIT, available from Invitrogen) allow one with ordinary skill in the art to deliver polynucleotides to target cells in culture and require minimal effort to optimize experimental parameters. In the alternative, transformation is performed using the calcium phosphate method

(Graham, F.L. and A.J. Eb (1973) Virology 52:456-467), or by electroporation (Neumann, E. et al. (1982) EMBO J. 1:841-845). The introduction of DNA to primary cells requires modification of these standardized mammalian transfection protocols.

In another embodiment of the invention, diseases or disorders caused by genetic defects with respect to FLEXHT expression are treated by constructing a retrovirus vector consisting of (i) the polynucleotide encoding FLEXHT under the control of an independent promoter or the retrovirus long terminal repeat (LTR) promoter, (ii) appropriate RNA packaging signals, and (iii) a Rev-responsive element (RRE) along with additional retrovirus *cis*-acting RNA sequences and coding sequences required for efficient vector propagation. Retrovirus vectors (e.g., PFB and PFBNEO) are commercially available (Stratagene) and are based on published data (Riviere, I. et al. (1995) Proc. Natl. Acad. Sci. U.S.A. 92:6733-6737), incorporated by reference herein. The vector is propagated in an appropriate vector producing cell line (VPCL) that expresses an envelope gene with a tropism for receptors on the target cells or a promiscuous envelope protein such as VSVg (Armentano, D. et al. (1987) J. Virol. 61:1647-1650; Bender, M.A. et al. (1987) J. Virol. 61:1639-1646; Adam, M.A. and A.D. Miller (1988) J. Virol. 62:3802-3806; Dull, T. et al. (1998) J. Virol. 72:8463-8471; Zufferey, R. et al. (1998) J. Virol. 72:9873-9880). U.S. Patent Number 5,910,434 to Rigg ("Method for obtaining retrovirus packaging cell lines producing high transducing efficiency retroviral supernatant") discloses a method for obtaining retrovirus packaging cell lines and is hereby incorporated by reference. Propagation of retrovirus vectors, transduction of a population of cells (e.g., $CD4^+$ T-cells), and the return of transduced cells to a patient are procedures well known to persons skilled in the art of gene therapy and have been well documented (Ranga, U. et al. (1997) J. Virol. 71:7020-7029; Bauer, G. et al. (1997) Blood 89:2259-2267; Bonyhadi, M.L. (1997) J. Virol. 71:4707-4716; Ranga, U. et al. (1998) Proc. Natl. Acad. Sci. U.S.A. 95:1201-1206; Su, L. (1997) Blood 89:2283-2290).

In the alternative, an adenovirus-based gene therapy delivery system is used to deliver polynucleotides encoding FLEXHT to cells which have one or more genetic abnormalities with respect to the expression of FLEXHT. The construction and packaging of adenovirus-based vectors are well known to those with ordinary skill in the art. Replication defective adenovirus vectors have proven to be versatile for importing genes encoding immunoregulatory proteins into intact islets in the pancreas (Csete, M.E. et al. (1995) Transplantation 27:263-268). Potentially useful adenoviral vectors are described in U.S. Patent Number 5,707,618 to Armentano ("Adenovirus vectors for gene therapy"), hereby incorporated by reference. For adenoviral vectors, see also Antinozzi, P.A. et al. (1999) Annu. Rev. Nutr. 19:511-544; and Verma, I.M. and N. Somia (1997) Nature 18:389:239-242, both incorporated by reference herein.

In another alternative, a herpes-based, gene therapy delivery system is used to deliver

polynucleotides encoding FLEXHT to target cells which have one or more genetic abnormalities with respect to the expression of FLEXHT. The use of herpes simplex virus (HSV)-based vectors may be especially valuable for introducing FLEXHT to cells of the central nervous system, for which HSV has a tropism. The construction and packaging of herpes-based vectors are well known to those with ordinary skill in the art. A replication-competent herpes simplex virus (HSV) type 1-based vector has been used to deliver a reporter gene to the eyes of primates (Liu, X. et al. (1999) Exp. Eye Res.169:385-395). The construction of a HSV-1 virus vector has also been disclosed in detail in U.S. Patent Number 5,804,413 to DeLuca ("Herpes simplex virus strains for gene transfer"), which is hereby incorporated by reference. U.S. Patent Number 5,804,413 teaches the use of recombinant HSV d92 which consists of a genome containing at least one exogenous gene to be transferred to a cell under the control of the appropriate promoter for purposes including human gene therapy. Also taught by this patent are the construction and use of recombinant HSV strains deleted for ICP4, ICP27 and ICP22. For HSV vectors, see also Goins, W.F. et al. (1999) J. Virol. 73:519-532 and Xu, H. et al. (1994) Dev. Biol. 163:152-161, hereby incorporated by reference. The manipulation of cloned herpesvirus sequences, the generation of recombinant virus following the transfection of multiple plasmids containing different segments of the large herpesvirus genomes, the growth and propagation of herpesvirus, and the infection of cells with herpesvirus are techniques well known to those of ordinary skill in the art.

In another alternative, an alphavirus (positive, single-stranded RNA virus) vector is used to deliver polynucleotides encoding FLEXHT to target cells. The biology of the prototypic alphavirus, Semliki Forest Virus (SFV), has been studied extensively and gene transfer vectors have been based on the SFV genome (Garoff, H. and K.-J. Li (1998) Curr. Opin. Biotech. 9:464-469). During alphavirus RNA replication, a subgenomic RNA is generated that normally encodes the viral capsid proteins. This subgenomic RNA replicates to higher levels than the full-length genomic RNA, resulting in the overproduction of capsid proteins relative to the viral proteins with enzymatic activity (e.g., protease and polymerase). Similarly, inserting the coding sequence for FLEXHT into the alphavirus genome in place of the capsid-coding region results in the production of a large number of FLEXHT-coding RNAs and the synthesis of high levels of FLEXHT in vector transduced cells. While alphavirus infection is typically associated with cell lysis within a few days, the ability to establish a persistent infection in hamster normal kidney cells (BHK-21) with a variant of Sindbis virus (SIN) indicates that the lytic replication of alphaviruses can be altered to suit the needs of the gene therapy application (Dryga, S.A. et al. (1997) Virology 228:74-83). The wide host range of alphaviruses will allow the introduction of FLEXHT into a variety of cell types. The specific transduction of a subset of cells in a population may require the sorting of cells prior to transduction. The methods of manipulating infectious cDNA clones

of alphaviruses, performing alphavirus cDNA and RNA transfections, and performing alphavirus infections, are well known to those with ordinary skill in the art.

Oligonucleotides derived from the transcription initiation site, e.g., between about positions -10 and +10 from the start site, may also be employed to inhibit gene expression. Similarly, inhibition can be achieved using triple helix base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature. (See, e.g., Gee, J.E. et al. (1994) in Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing, Mt. Kisco NY, pp. 163-177.) A complementary sequence or antisense molecule may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. For example, engineered hammerhead motif ribozyme molecules may specifically and efficiently catalyze endonucleolytic cleavage of sequences encoding FLEXHT.

Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites, including the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides, corresponding to the region of the target gene containing the cleavage site, may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

Complementary ribonucleic acid molecules and ribozymes of the invention may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by in vitro and in vivo transcription of DNA sequences encoding FLEXHT. Such DNA sequences may be incorporated into a wide variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize complementary RNA, constitutively or inducibly, can be introduced into cell lines, cells, or tissues.

RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule, or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages

within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of nontraditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

An additional embodiment of the invention encompasses a method for screening for a compound which is effective in altering expression of a polynucleotide encoding FLEXHT. Compounds which may be effective in altering expression of a specific polynucleotide may include, but are not limited to, oligonucleotides, antisense oligonucleotides, triple helix-forming oligonucleotides, transcription factors and other polypeptide transcriptional regulators, and non-macromolecular chemical entities which are capable of interacting with specific polynucleotide sequences. Effective compounds may alter polynucleotide expression by acting as either inhibitors or promoters of polynucleotide expression. Thus, in the treatment of disorders associated with increased FLEXHT expression or activity, a compound which specifically inhibits expression of the polynucleotide encoding FLEXHT may be therapeutically useful, and in the treament of disorders associated with decreased FLEXHT expression or activity, a compound which specifically promotes expression of the polynucleotide encoding FLEXHT may be therapeutically useful.

At least one, and up to a plurality, of test compounds may be screened for effectiveness in altering expression of a specific polynucleotide. A test compound may be obtained by any method commonly known in the art, including chemical modification of a compound known to be effective in altering polynucleotide expression; selection from an existing, commercially-available or proprietary library of naturally-occurring or non-natural chemical compounds; rational design of a compound based on chemical and/or structural properties of the target polynucleotide; and selection from a library of chemical compounds created combinatorially or randomly. A sample comprising a polynucleotide encoding FLEXHT is exposed to at least one test compound thus obtained. The sample may comprise, for example, an intact or permeabilized cell, or an in vitro cell-free or reconstituted biochemical system. Alterations in the expression of a polynucleotide encoding FLEXHT are assayed by any method commonly known in the art. Typically, the expression of a specific nucleotide is detected by hybridization with a probe having a nucleotide sequence complementary to the sequence of the polynucleotide encoding FLEXHT. The amount of hybridization may be quantified, thus forming the basis for a comparison of the expression of the polynucleotide both with and without exposure to one or more test compounds. Detection of a change in the expression of a polynucleotide exposed to a test compound indicates that the test compound is effective in altering the expression of the polynucleotide. A screen for a compound effective in altering expression of a specific polynucleotide can be carried out, for example, using a Schizosaccharomyces pombe gene expression system (Atkins, D. et al. (1999) U.S. Patent No.

5,932,435; Arndt, G.M. et al. (2000) Nucleic Acids Res. 28:E15) or a human cell line such as HeLa cell (Clarke, M.L. et al. (2000) Biochem. Biophys. Res. Commun. 268:8-13). A particular embodiment of the present invention involves screening a combinatorial library of oligonucleotides (such as deoxyribonucleotides, ribonucleotides, peptide nucleic acids, and modified oligonucleotides) for antisense activity against a specific polynucleotide sequence (Bruice, T.W. et al. (1997) U.S. Patent No. 5,686,242; Bruice, T.W. et al. (2000) U.S. Patent No. 6,022,691).

Many methods for introducing vectors into cells or tissues are available and equally suitable for use in vivo, in vitro, and ex vivo. For ex vivo therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection, by liposome injections, or by polycationic amino polymers may be achieved using methods which are well known in the art. (See, e.g., Goldman, C.K. et al. (1997) Nat. Biotechnol. 15:462-466.)

Any of the therapeutic methods described above may be applied to any subject in need of such therapy, including, for example, mammals such as humans, dogs, cats, cows, horses, rabbits, and monkeys.

An additional embodiment of the invention relates to the administration of a pharmaceutical composition which generally comprises an active ingredient formulated with a pharmaceutically acceptable excipient. Excipients may include, for example, sugars, starches, celluloses, gums, and proteins. Various formulations are commonly known and are thoroughly discussed in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing, Easton PA). Such pharmaceutical compositions may consist of FLEXHT, antibodies to FLEXHT, and mimetics, agonists, antagonists, or inhibitors of FLEXHT.

The pharmaceutical compositions utilized in this invention may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, pulmonary, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal means.

Pharmaceutical compositions for pulmonary administration may be prepared in liquid or dry powder form. These compositions are generally aerosolized immediately prior to inhalation by the patient. In the case of small molecules (e.g. traditional low molecular weight organic drugs), aerosol delivery of fast-acting formulations is well-known in the art. In the case of macromolecules (e.g. larger peptides and proteins), recent developments in the field of pulmonary delivery via the alveolar region of the lung have enabled the practical delivery of drugs such as insulin to blood circulation (see, e.g., Patton, J.S. et al., U.S. Patent No. 5,997,848). Pulmonary delivery has the advantage of administration without needle injection, and obviates the need for potentially toxic penetration enhancers.

Pharmaceutical compositions suitable for use in the invention include compositions wherein the

active ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art.

Specialized forms of pharmaceutical compositions may be prepared for direct intracellular delivery of macromolecules comprising FLEXHT or fragments thereof. For example, liposome preparations containing a cell-impermeable macromolecule may promote cell fusion and intracellular delivery of the macromolecule. Alternatively, FLEXHT or a fragment thereof may be joined to a short cationic N-terminal portion from the HIV Tat-1 protein. Fusion proteins thus generated have been found to transduce into the cells of all tissues, including the brain, in a mouse model system (Schwarze, S.R. et al. (1999) Science 285:1569-1572).

For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays, e.g., of neoplastic cells, or in animal models such as mice, rats, rabbits, dogs, monkeys, or pigs. An animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

A therapeutically effective dose refers to that amount of active ingredient, for example FLEXHT or fragments thereof, antibodies of FLEXHT, and agonists, antagonists or inhibitors of FLEXHT, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or with experimental animals, such as by calculating the ED_{50} (the dose therapeutically effective in 50% of the population) or LD_{50} (the dose lethal to 50% of the population) statistics. The dose ratio of toxic to therapeutic effects is the therapeutic index, which can be expressed as the LD_{50}/ED_{50} ratio. Pharmaceutical compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are used to formulate a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that includes the ED_{50} with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, the sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the subject requiring treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors which may be taken into account include the severity of the disease state, the general health of the subject, the age, weight, and gender of the subject, time and frequency of administration, drug combination(s), reaction sensitivities, and response to therapy. Long-acting pharmaceutical compositions may be administered every 3 to 4 days, every week, or biweekly depending on the half-life and clearance rate of the particular formulation.

Normal dosage amounts may vary from about 0.1 μg to 100,000 μg, up to a total dose of

about 1 gram, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

DIAGNOSTICS

In another embodiment, antibodies which specifically bind FLEXHT may be used for the diagnosis of disorders characterized by expression of FLEXHT, or in assays to monitor patients being treated with FLEXHT or agonists, antagonists, or inhibitors of FLEXHT. Antibodies useful for diagnostic purposes may be prepared in the same manner as described above for therapeutics. Diagnostic assays for FLEXHT include methods which utilize the antibody and a label to detect FLEXHT in human body fluids or in extracts of cells or tissues. The antibodies may be used with or without modification, and may be labeled by covalent or non-covalent attachment of a reporter molecule. A wide variety of reporter molecules, several of which are described above, are known in the art and may be used.

A variety of protocols for measuring FLEXHT, including ELISAs, RIAs, and FACS, are known in the art and provide a basis for diagnosing altered or abnormal levels of FLEXHT expression. Normal or standard values for FLEXHT expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, for example, human subjects, with antibody to FLEXHT under conditions suitable for complex formation. The amount of standard complex formation may be quantitated by various methods, such as photometric means. Quantities of FLEXHT expressed in subject, control, and disease samples from biopsied tissues are compared with the standard values. Deviation between standard and subject values establishes the parameters for diagnosing disease.

In another embodiment of the invention, the polynucleotides encoding FLEXHT may be used for diagnostic purposes. The polynucleotides which may be used include oligonucleotide sequences, complementary RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect and quantify gene expression in biopsied tissues in which expression of FLEXHT may be correlated with disease. The diagnostic assay may be used to determine absence, presence, and excess expression of FLEXHT, and to monitor regulation of FLEXHT levels during therapeutic intervention.

In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding FLEXHT or closely related molecules may be used to identify nucleic acid sequences which encode FLEXHT. The specificity of the probe, whether it is made from a highly specific region, e.g., the 5' regulatory region, or from a less specific region, e.g., a conserved motif, and the stringency of the hybridization or amplification will determine whether the

probe identifies only naturally occurring sequences encoding FLEXHT, allelic variants, or related sequences.

Probes may also be used for the detection of related sequences, and may have at least 50% sequence identity to any of the FLEXHT encoding sequences. The hybridization probes of the subject invention may be DNA or RNA and may be derived from the sequence of SEQ ID NO:56-110 or from genomic sequences including promoters, enhancers, and introns of the FLEXHT gene.

Means for producing specific hybridization probes for DNAs encoding FLEXHT include the cloning of polynucleotide sequences encoding FLEXHT or FLEXHT derivatives into vectors for the production of mRNA probes. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by means of the addition of the appropriate RNA polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a variety of reporter groups, for example, by radionuclides such as ^{32}P or ^{35}S, or by enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

Polynucleotide sequences encoding FLEXHT may be used for the diagnosis of disorders associated with expression of FLEXHT. Examples of such disorders include, but are not limited to, a developmental disorder such as renal tubular acidosis, anemia, Cushing's syndrome, achondroplastic dwarfism, Duchenne and Becker muscular dystrophy, epilepsy, gonadal dysgenesis, WAGR syndrome (Wilms' tumor, aniridia, genitourinary abnormalities, and mental retardation), Smith-Magenis syndrome, myelodysplastic syndrome, hereditary mucoepithelial dysplasia, hereditary keratodermas, hereditary neuropathies such as Charcot-Marie-Tooth disease and neurofibromatosis, hypothyroidism, hydrocephalus, a seizure disorder such as Syndenham's chorea and cerebral palsy, spina bifida, anencephaly, craniorachischisis, congenital glaucoma, cataract, and sensorineural hearing loss; a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, a cancer of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; and an immunological disorder such as inflammation, actinic keratosis, acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, arteriosclerosis, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, bronchitis, bursitis, cholecystitis, cirrhosis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, erythroblastosis fetalis, erythema nodosum,

atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, paroxysmal nocturnal hemoglobinuria, hepatitis, hypereosinophilia, irritable bowel syndrome, episodic lymphopenia with lymphocytotoxins, mixed connective tissue disease (MCTD), multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, myelofibrosis, osteoarthritis, osteoporosis, pancreatitis, polycythemia vera, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, primary thrombocythemia, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, trauma, and hematopoietic cancer including lymphoma, leukemia, and myeloma. The polynucleotide sequences encoding FLEXHT may be used in Southern or northern analysis, dot blot, or other membrane-based technologies; in PCR technologies; in dipstick, pin, and multiformat ELISA-like assays; and in microarrays utilizing fluids or tissues from patients to detect altered FLEXHT expression. Such qualitative or quantitative methods are well known in the art.

In a particular aspect, the nucleotide sequences encoding FLEXHT may be useful in assays that detect the presence of associated disorders, particularly those mentioned above. The nucleotide sequences encoding FLEXHT may be labeled by standard methods and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantified and compared with a standard value. If the amount of signal in the patient sample is significantly altered in comparison to a control sample then the presence of altered levels of nucleotide sequences encoding FLEXHT in the sample indicates the presence of the associated disorder. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or to monitor the treatment of an individual patient.

In order to provide a basis for the diagnosis of a disorder associated with expression of FLEXHT, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence, or a fragment thereof, encoding FLEXHT, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with values from an experiment in which a known amount of a substantially purified polynucleotide is used. Standard values obtained in this manner may be compared with values obtained from samples from patients who are symptomatic for a disorder. Deviation from standard values is used to establish the presence of a disorder.

Once the presence of a disorder is established and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to determine if the level of expression in the

patient begins to approximate that which is observed in the normal subject. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

With respect to cancer, the presence of an abnormal amount of transcript (either under- or overexpressed) in biopsied tissue from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby preventing the development or further progression of the cancer.

Additional diagnostic uses for oligonucleotides designed from the sequences encoding FLEXHT may involve the use of PCR. These oligomers may be chemically synthesized, generated enzymatically, or produced in vitro. Oligomers will preferably contain a fragment of a polynucleotide encoding FLEXHT, or a fragment of a polynucleotide complementary to the polynucleotide encoding FLEXHT, and will be employed under optimized conditions for identification of a specific gene or condition. Oligomers may also be employed under less stringent conditions for detection or quantification of closely related DNA or RNA sequences.

In a particular aspect, oligonucleotide primers derived from the polynucleotide sequences encoding FLEXHT may be used to detect single nucleotide polymorphisms (SNPs). SNPs are substitutions, insertions and deletions that are a frequent cause of inherited or acquired genetic disease in humans. Methods of SNP detection include, but are not limited to, single-stranded conformation polymorphism (SSCP) and fluorescent SSCP (fSSCP) methods. In SSCP, oligonucleotide primers derived from the polynucleotide sequences encoding FLEXHT are used to amplify DNA using the polymerase chain reaction (PCR). The DNA may be derived, for example, from diseased or normal tissue, biopsy samples, bodily fluids, and the like. SNPs in the DNA cause differences in the secondary and tertiary structures of PCR products in single-stranded form, and these differences are detectable using gel electrophoresis in non-denaturing gels. In fSCCP, the oligonucleotide primers are fluorescently labeled, which allows detection of the amplimers in high-throughput equipment such as DNA sequencing machines. Additionally, sequence database analysis methods, termed in silico SNP (isSNP), are capable of identifying polymorphisms by comparing the sequence of individual overlapping DNA fragments which assemble into a common consensus sequence. These computer-based methods filter out sequence variations due to laboratory preparation of DNA and sequencing errors using statistical models and automated analyses of DNA sequence chromatograms. In the alternative, SNPs may be detected and characterized by mass spectrometry using, for example, the high throughput MASSARRAY system (Sequenom, Inc., San Diego CA).

Methods which may also be used to quantify the expression of FLEXHT include radiolabeling or biotinylating nucleotides, coamplification of a control nucleic acid, and interpolating results from standard curves. (See, e.g., Melby, P.C. et al. (1993) J. Immunol. Methods 159:235-244; Duplaa, C. et al. (1993) Anal. Biochem. 212:229-236.) The speed of quantitation of multiple samples may be accelerated by running the assay in a high-throughput format where the oligomer or polynucleotide of interest is presented in various dilutions and a spectrophotometric or colorimetric response gives rapid quantitation.

The polynucleotides are also useful for identifying individuals on the basis of minute biological samples, for example, by matching the restriction fragment length polymorphism (RFLP) pattern of a sample's DNA to that of an individual's DNA. The polynucleotides of the present invention can also be used to determine the actual base-by-base DNA sequence of selected portions of an individual's genome. These sequences can be used to prepare PCR primers for amplifying and isolating such selected DNA, which can then be sequenced. Using this technique, an individual can be identified through a unique set of DNA sequences. Once a unique ID database is established for an individual, positive identification of that individual can be made from extremely small tissue samples.

DNA-based identification techniques are critical in forensic technology. DNA sequences taken from very small biological samples such as tissues, e.g., hair, skin, or body fluids (e.g., blood, saliva, semen, etc.), can be amplified using, e.g., PCR, to identify individuals. (See, e.g., Erlich, H. (1992) PCR Technology, Freeman and Co., New York NY). Similarly, polynucleotides of the present invention can be used as polymorphic markers.

There is also a need for reagents capable of identifying the source of a particular tissue. Appropriate reagents can comprise, for example, DNA probes or primers prepared from the sequences of the present invention that are specific for particular tissues. Panels of such reagents can identify tissue by species and/or by organ type. In a similar fashion, these reagents can be used to screen tissue cultures for contamination.

The polynucleotides of the present invention can also be used as molecular weight markers on nucleic acid gels or Southern blots, as diagnostic probes for the presence of a specific mRNA in a particular cell type, in the creation of subtracted cDNA libraries which aid in the discovery of novel polynucleotides, in selection and synthesis of oligomers for attachment to an array or other support, and as an antigen to elicit an immune response.

In further embodiments, oligonucleotides or longer fragments derived from any of the polynucleotide sequences described herein may be used as elements on a microarray. The microarray can be used in transcript imaging techniques which monitor the relative expression levels of large numbers of genes simultaneously as described in Seilhamer, J.J. et al., "Comparative Gene Transcript Analysis," U.S. Patent No. 5,840,484, incorporated herein by reference. The microarray may also be

used to identify genetic variants, mutations, and polymorphisms. This information may be used to determine gene function, to understand the genetic basis of a disorder, to diagnose a disorder, to monitor progression/regression of disease as a function of gene expression, and to develop and monitor the activities of therapeutic agents in the treatment of disease. In particular, this information may be used to develop a pharmacogenomic profile of a patient in order to select the most appropriate and effective treatment regimen for that patient. For example, therapeutic agents which are highly effective and display the fewest side effects may be selected for a patient based on his/her pharmacogenomic profile.

In another embodiment, antibodies specific for FLEXHT, or FLEXHT or fragments thereof may be used as elements on a microarray. The microarray may be used to monitor or measure protein-protein interactions, drug-target interactions, and gene expression profiles, as described above.

Microarrays may be prepared, used, and analyzed using methods known in the art. (See, e.g., Brennan, T.M. et al. (1995) U.S. Patent No. 5,474,796; Schena, M. et al. (1996) Proc. Natl. Acad. Sci. USA 93:10614-10619; Baldeschweiler et al. (1995) PCT application WO95/251116; Shalon, D. et al. (1995) PCT application WO95/35505; Heller, R.A. et al. (1997) Proc. Natl. Acad. Sci. USA 94:2150-2155; and Heller, M.J. et al. (1997) U.S. Patent No. 5,605,662.) Various types of microarrays are well known and thoroughly described in DNA Microarrays: A Practical Approach, M. Schena, ed. (1999) Oxford University Press, London, hereby expressly incorporated by reference.

In another embodiment of the invention, nucleic acid sequences encoding FLEXHT may be used to generate hybridization probes useful in mapping the naturally occurring genomic sequence. Either coding or noncoding sequences may be used, and in some instances, noncoding sequences may be preferable over coding sequences. For example, conservation of a coding sequence among members of a multi-gene family may potentially cause undesired cross hybridization during chromosomal mapping. The sequences may be mapped to a particular chromosome, to a specific region of a chromosome, or to artificial chromosome constructions, e.g., human artificial chromosomes (HACs), yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial P1 constructions, or single chromosome cDNA libraries. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355; Price, C.M. (1993) Blood Rev. 7:127-134; and Trask, B.J. (1991) Trends Genet. 7:149-154.) Once mapped, the nucleic acid sequences of the invention may be used to develop genetic linkage maps, for example, which correlate the inheritance of a disease state with the inheritance of a particular chromosome region or restriction fragment length polymorphism (RFLP). (See, e.g., Lander, E.S. and D. Botstein (1986) Proc. Natl. Acad. Sci. USA 83:7353-7357.)

Fluorescent in situ hybridization (FISH) may be correlated with other physical and genetic map data. (See, e.g., Heinz-Ulrich, et al. (1995) in Meyers, supra, pp. 965-968.) Examples of genetic map data can be found in various scientific journals or at the Online Mendelian Inheritance in Man (OMIM) World Wide Web site. Correlation between the location of the gene encoding FLEXHT on a physical

map and a specific disorder, or a predisposition to a specific disorder, may help define the region of DNA associated with that disorder and thus may further positional cloning efforts.

In situ hybridization of chromosomal preparations and physical mapping techniques, such as linkage analysis using established chromosomal markers, may be used for extending genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers even if the exact chromosomal locus is not known. This information is valuable to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once the gene or genes responsible for a disease or syndrome have been crudely localized by genetic linkage to a particular genomic region, e.g., ataxia-telangiectasia to 11q22-23, any sequences mapping to that area may represent associated or regulatory genes for further investigation. (See, e.g., Gatti, R.A. et al. (1988) Nature 336:577-580.) The nucleotide sequence of the instant invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc., among normal, carrier, or affected individuals.

In another embodiment of the invention, FLEXHT, its catalytic or immunogenic fragments, or oligopeptides thereof can be used for screening libraries of compounds in any of a variety of drug screening techniques. The fragment employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes between FLEXHT and the agent being tested may be measured.

Another technique for drug screening provides for high throughput screening of compounds having suitable binding affinity to the protein of interest. (See, e.g., Geysen, et al. (1984) PCT application WO84/03564.) In this method, large numbers of different small test compounds are synthesized on a solid substrate. The test compounds are reacted with FLEXHT, or fragments thereof, and washed. Bound FLEXHT is then detected by methods well known in the art. Purified FLEXHT can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

In another embodiment, one may use competitive drug screening assays in which neutralizing antibodies capable of binding FLEXHT specifically compete with a test compound for binding FLEXHT. In this manner, antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with FLEXHT.

In additional embodiments, the nucleotide sequences which encode FLEXHT may be used in any molecular biology techniques that have yet to be developed, provided the new techniques rely on properties of nucleotide sequences that are currently known, including, but not limited to, such properties as the triplet genetic code and specific base pair interactions.

Without further elaboration, it is believed that one skilled in the art can, using the preceding description, utilize the present invention to its fullest extent. The following preferred specific embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

The disclosures of all patents, applications, and publications, mentioned above and below, in particular U.S. Ser. No. 09/311,894, U.S. Ser. No. 09/311,940, and U.S. Ser. No. 09/311,937, are hereby expressly incorporated by reference.

EXAMPLES

I. Construction of cDNA Libraries

RNA was purchased from Clontech or isolated from tissues described in Table 4. Some tissues were homogenized and lysed in guanidinium isothiocyanate, while others were homogenized and lysed in phenol or in a suitable mixture of denaturants, such as TRIZOL (Life Technologies), a monophasic solution of phenol and guanidine isothiocyanate. The resulting lysates were centrifuged over CsCl cushions or extracted with chloroform. RNA was precipitated from the lysates with either isopropanol or sodium acetate and ethanol, or by other routine methods.

Phenol extraction and precipitation of RNA were repeated as necessary to increase RNA purity. In some cases, RNA was treated with DNase. For most libraries, poly(A+) RNA was isolated using oligo d(T)-coupled paramagnetic particles (Promega), OLIGOTEX latex particles (QIAGEN, Chatsworth CA), or an OLIGOTEX mRNA purification kit (QIAGEN). Alternatively, RNA was isolated directly from tissue lysates using other RNA isolation kits, e.g., the POLY(A)PURE mRNA purification kit (Ambion, Austin TX).

In some cases, Stratagene was provided with RNA and constructed the corresponding cDNA libraries. Otherwise, cDNA was synthesized and cDNA libraries were constructed with the UNIZAP vector system (Stratagene) or SUPERSCRIPT plasmid system (Life Technologies), using the recommended procedures or similar methods known in the art. (See, e.g., Ausubel, 1997, supra, units 5.1-6.6.) Reverse transcription was initiated using oligo d(T) or random primers. Synthetic oligonucleotide adapters were ligated to double stranded cDNA, and the cDNA was digested with the appropriate restriction enzyme or enzymes. For most libraries, the cDNA was size-selected (300-1000 bp) using SEPHACRYL S1000, SEPHAROSE CL2B, or SEPHAROSE CL4B column chromatography (Amersham Pharmacia Biotech) or preparative agarose gel electrophoresis. cDNAs were ligated into compatible restriction enzyme sites of the polylinker of a suitable plasmid, e.g., PBLUESCRIPT plasmid (Stratagene), PSPORT1 plasmid (Life Technologies), pcDNA2.1 plasmid (Invitrogen, Carlsbad CA), or pINCY plasmid (Incyte Genomics, Palo Alto CA). Recombinant plasmids were transformed into competent E. coli cells including XL1-Blue, XL1-BlueMRF, or SOLR

from Stratagene or DH5α, DH10B, or ElectroMAX DH10B from Life Technologies.

II. Isolation of cDNA Clones

Plasmids obtained as described in Example I were recovered from host cells by in vivo excision using the UNIZAP vector system (Stratagene) or by cell lysis. Plasmids were purified using at least one of the following: a Magic or WIZARD Minipreps DNA purification system (Promega); an AGTC Miniprep purification kit (Edge Biosystems, Gaithersburg MD); and QIAWELL 8 Plasmid, QIAWELL 8 Plus Plasmid, QIAWELL 8 Ultra Plasmid purification systems or the R.E.A.L. PREP 96 plasmid purification kit from QIAGEN. Following precipitation, plasmids were resuspended in 0.1 ml of distilled water and stored, with or without lyophilization, at 4°C.

Alternatively, plasmid DNA was amplified from host cell lysates using direct link PCR in a high-throughput format (Rao, V.B. (1994) Anal. Biochem. 216:1-14). Host cell lysis and thermal cycling steps were carried out in a single reaction mixture. Samples were processed and stored in 384-well plates, and the concentration of amplified plasmid DNA was quantified fluorometrically using PICOGREEN dye (Molecular Probes, Eugene OR) and a FLUOROSKAN II fluorescence scanner (Labsystems Oy, Helsinki, Finland).

III. Sequencing and Analysis

Incyte cDNA recovered in plasmids as described in Example II were sequenced as follows. Sequencing reactions were processed using standard methods or high-throughput instrumentation such as the ABI CATALYST 800 (Perkin-Elmer) thermal cycler or the PTC-200 thermal cycler (MJ Research) in conjunction with the HYDRA microdispenser (Robbins Scientific) or the MICROLAB 2200 (Hamilton) liquid transfer system. cDNA sequencing reactions were prepared using reagents provided by Amersham Pharmacia Biotech or supplied in ABI sequencing kits such as the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Perkin-Elmer). Electrophoretic separation of cDNA sequencing reactions and detection of labeled polynucleotides were carried out using the MEGABACE 1000 DNA sequencing system (Molecular Dynamics); the ABI PRISM 373 or 377 sequencing system (Perkin-Elmer) in conjunction with standard ABI protocols and base calling software; or other sequence analysis systems known in the art. Reading frames within the cDNA sequences were identified using standard methods (reviewed in Ausubel, 1997, supra, unit 7.7). Some of the cDNA sequences were selected for extension using the techniques disclosed in Example V.

The polynucleotide sequences derived from cDNA sequencing were assembled and analyzed using a combination of software programs which utilize algorithms well known to those skilled in the art. Table 6 summarizes the tools, programs, and algorithms used and provides applicable descriptions, references, and threshold parameters. The first column of Table 6 shows the tools, programs, and algorithms used, the second column provides brief descriptions thereof, the third column presents

appropriate references, all of which are incorporated by reference herein in their entirety, and the fourth column presents, where applicable, the scores, probability values, and other parameters used to evaluate the strength of a match between two sequences (the higher the score, the greater the homology between two sequences). Sequences were analyzed using MACDNASIS PRO software (Hitachi Software Engineering, South San Francisco CA) and LASERGENE software (DNASTAR). Polynucleotide and polypeptide sequence alignments were generated using the default parameters specified by the clustal algorithm as incorporated into the MEGALIGN multisequence alignment program (DNASTAR), which also calculates the percent identity between aligned sequences.

The polynucleotide sequences were validated by removing vector, linker, and polyA sequences and by masking ambiguous bases, using algorithms and programs based on BLAST, dynamic programing, and dinucleotide nearest neighbor analysis. The sequences were then queried against a selection of public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases, and BLOCKS, PRINTS, DOMO, PRODOM, and PFAM to acquire annotation using programs based on BLAST, FASTA, and BLIMPS. The sequences were assembled into full length polynucleotide sequences using programs based on Phred, Phrap, and Consed, and were screened for open reading frames using programs based on GeneMark, BLAST, and FASTA. The full length polynucleotide sequences were translated to derive the corresponding full length amino acid sequences, and these full length sequences were subsequently analyzed by querying against databases such as the GenBank databases (described above), SwissProt, BLOCKS, PRINTS, DOMO, PRODOM, Prosite, and Hidden Markov Model (HMM)-based protein family databases such as PFAM. HMM is a probabilistic approach which analyzes consensus primary structures of gene families. (See, e.g., Eddy, S.R. (1996) Curr. Opin. Struct. Biol. 6:361-365.)

The programs described above for the assembly and analysis of full length polynucleotide and amino acid sequences were also used to identify polynucleotide sequence fragments from SEQ ID NO:56-110. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and amplification technologies were described in The Invention section above.

IV. Analysis of Polynucleotide Expression

Northern analysis is a laboratory technique used to detect the presence of a transcript of a gene and involves the hybridization of a labeled nucleotide sequence to a membrane on which RNAs from a particular cell type or tissue have been bound. (See, e.g., Sambrook, supra, ch. 7; Ausubel, 1995, supra, ch. 4 and 16.)

Analogous computer techniques applying BLAST were used to search for identical or related molecules in cDNA databases such as GenBank or LIFESEQ (Incyte Genomics). This analysis is much faster than multiple membrane-based hybridizations. In addition, the sensitivity of the computer

search can be modified to determine whether any particular match is categorized as exact or similar. The basis of the search is the product score, which is defined as:

$$\frac{\text{BLAST Score x Percent Identity}}{\text{5 x minimum \{length(Seq. 1), length(Seq. 2)\}}}$$

The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. The product score is a normalized value between 0 and 100, and is calculated as follows: the BLAST score is multiplied by the percent nucleotide identity and the product is divided by (5 times the length of the shorter of the two sequences). The BLAST score is calculated by assigning a score of +5 for every base that matches in a high-scoring segment pair (HSP), and -4 for every mismatch. Two sequences may share more than one HSP (separated by gaps). If there is more than one HSP, then the pair with the highest BLAST score is used to calculate the product score. The product score represents a balance between fractional overlap and quality in a BLAST alignment. For example, a product score of 100 is produced only for 100% identity over the entire length of the shorter of the two sequences being compared. A product score of 70 is produced either by 100% identity and 70% overlap at one end, or by 88% identity and 100% overlap at the other. A product score of 50 is produced either by 100% identity and 50% overlap at one end, or 79% identity and 100% overlap.

The results of northern analyses are reported as a percentage distribution of libraries in which the transcript encoding FLEXHT occurred. Analysis involved the categorization of cDNA libraries by organ/tissue and disease. The organ/tissue categories included cardiovascular, dermatologic, developmental, endocrine, gastrointestinal, hematopoietic/immune, musculoskeletal, nervous, reproductive, and urologic. The disease/condition categories included cancer, inflammation, trauma, cell proliferation, neurological, and pooled. For each category, the number of libraries expressing the sequence of interest was counted and divided by the total number of libraries across all categories. Percentage values of tissue-specific and disease- or condition-specific expression are reported in Table 3.

V. Extension of FLEXHT Encoding Polynucleotides

The full length nucleic acid sequences of SEQ ID NO:56-110 were produced by extension of an appropriate fragment of the full length molecule using oligonucleotide primers designed from this fragment. One primer was synthesized to initiate 5' extension of the known fragment, and the other primer, to initiate 3' extension of the known fragment. The initial primers were designed using OLIGO 4.06 software (National Biosciences), or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the target sequence at

temperatures of about 68°C to about 72°C. Any stretch of nucleotides which would result in hairpin structures and primer-primer dimerizations was avoided.

Selected human cDNA libraries were used to extend the sequence. If more than one extension was necessary or desired, additional or nested sets of primers were designed.

High fidelity amplification was obtained by PCR using methods well known in the art. PCR was performed in 96-well plates using the PTC-200 thermal cycler (MJ Research, Inc.). The reaction mix contained DNA template, 200 nmol of each primer, reaction buffer containing Mg^{2+}, $(NH_4)_2SO_4$, and β-mercaptoethanol, Taq DNA polymerase (Amersham Pharmacia Biotech), ELONGASE enzyme (Life Technologies), and Pfu DNA polymerase (Stratagene), with the following parameters for primer pair PCI A and PCI B: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C. In the alternative, the parameters for primer pair T7 and SK+ were as follows: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 57°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C.

The concentration of DNA in each well was determined by dispensing 100 μl PICOGREEN quantitation reagent (0.25% (v/v) PICOGREEN; Molecular Probes, Eugene OR) dissolved in 1X TE and 0.5 μl of undiluted PCR product into each well of an opaque fluorimeter plate (Corning Costar, Acton MA), allowing the DNA to bind to the reagent. The plate was scanned in a Fluoroskan II (Labsystems Oy, Helsinki, Finland) to measure the fluorescence of the sample and to quantify the concentration of DNA. A 5 μl to 10 μl aliquot of the reaction mixture was analyzed by electrophoresis on a 1 % agarose mini-gel to determine which reactions were successful in extending the sequence.

The extended nucleotides were desalted and concentrated, transferred to 384-well plates, digested with CviJI cholera virus endonuclease (Molecular Biology Research, Madison WI), and sonicated or sheared prior to religation into pUC 18 vector (Amersham Pharmacia Biotech). For shotgun sequencing, the digested nucleotides were separated on low concentration (0.6 to 0.8%) agarose gels, fragments were excised, and agar digested with Agar ACE (Promega). Extended clones were religated using T4 ligase (New England Biolabs, Beverly MA) into pUC 18 vector (Amersham Pharmacia Biotech), treated with Pfu DNA polymerase (Stratagene) to fill-in restriction site overhangs, and transfected into competent E. coli cells. Transformed cells were selected on antibiotic-containing media, and individual colonies were picked and cultured overnight at 37°C in 384-well plates in LB/2x carb liquid media.

The cells were lysed, and DNA was amplified by PCR using Taq DNA polymerase (Amersham Pharmacia Biotech) and Pfu DNA polymerase (Stratagene) with the following parameters: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 72°C, 2 min; Step 5: steps 2,

3, and 4 repeated 29 times; Step 6: 72°C, 5 min; Step 7: storage at 4°C. DNA was quantified by PICOGREEN reagent (Molecular Probes) as described above. Samples with low DNA recoveries were reamplified using the same conditions as described above. Samples were diluted with 20% dimethysulfoxide (1:2, v/v), and sequenced using DYENAMIC energy transfer sequencing primers and the DYENAMIC DIRECT kit (Amersham Pharmacia Biotech) or the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Perkin-Elmer).

In like manner, the polynucleotide sequences of SEQ ID NO:56-110 are used to obtain 5' regulatory sequences using the procedure above, along with oligonucleotides designed for such extension, and an appropriate genomic library.

VI. Labeling and Use of Individual Hybridization Probes

Hybridization probes derived from SEQ ID NO:56-110 are employed to screen cDNAs, genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base pairs, is specifically described, essentially the same procedure is used with larger nucleotide fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 software (National Biosciences) and labeled by combining 50 pmol of each oligomer, 250 μCi of [γ-^{32}P] adenosine triphosphate (Amersham Pharmacia Biotech), and T4 polynucleotide kinase (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified using a SEPHADEX G-25 superfine size exclusion dextran bead column (Amersham Pharmacia Biotech). An aliquot containing 10^7 counts per minute of the labeled probe is used in a typical membrane-based hybridization analysis of human genomic DNA digested with one of the following endonucleases: Ase I, Bgl II, Eco RI, Pst I, Xba I, or Pvu II (DuPont NEN).

The DNA from each digest is fractionated on a 0.7% agarose gel and transferred to nylon membranes (Nytran Plus, Schleicher & Schuell, Durham NH). Hybridization is carried out for 16 hours at 40°C. To remove nonspecific signals, blots are sequentially washed at room temperature under conditions of up to, for example, 0.1 x saline sodium citrate and 0.5% sodium dodecyl sulfate. Hybridization patterns are visualized using autoradiography or an alternative imaging means and compared.

VII. Transcript Image Analysis

Transcript images are generated as described in Seilhamer, J.J. et al., "Comparative Gene Transcript Analysis," U.S. Patent No. 5,840,484, incorporated herein by reference.

VIII. Microarrays

The linkage or synthesis of array elements upon a microarray can be achieved utilizing photolithography, piezoelectric printing (ink-jet printing, See, e.g., Baldeschweiler, supra), mechanical microspotting technologies, and derivatives thereof. The substrate in each of the aforementioned

technologies should be uniform and solid with a non-porous surface (Schena (1999), supra). Suggested substrates include silicon, silica, glass slides, glass chips, and silicon wafers. Alternatively, a procedure analogous to a dot or slot blot may also be used to arrange and link elements to the surface of a substrate using thermal, UV, chemical, or mechanical bonding procedures. A typical array may be produced using available methods and machines well known to those of ordinary skill in the art and may contain any appropriate number of elements. (See, e.g., Schena, M. et al. (1995) Science 270:467-470; Shalon, D. et al. (1996) Genome Res. 6:639-645; Marshall, A. and J. Hodgson (1998) Nat. Biotechnol. 16:27-31.)

Full length cDNAs, Expressed Sequence Tags (ESTs), or fragments or oligomers thereof may comprise the elements of the microarray. Fragments or oligomers suitable for hybridization can be selected using software well known in the art such as LASERGENE software (DNASTAR). The array elements are hybridized with polynucleotides in a biological sample. The polynucleotides in the biological sample are conjugated to a fluorescent label or other molecular tag for ease of detection. After hybridization, nonhybridized nucleotides from the biological sample are removed, and a fluorescence scanner is used to detect hybridization at each array element. Alternatively, laser desorbtion and mass spectrometry may be used for detection of hybridization. The degree of complementarity and the relative abundance of each polynucleotide which hybridizes to an element on the microarray may be assessed. In one embodiment, microarray preparation and usage is described in detail below.

Tissue or Cell Sample Preparation

Total RNA is isolated from tissue samples using the guanidinium thiocyanate method and poly(A)$^+$ RNA is purified using the oligo-(dT) cellulose method. Each poly(A)$^+$ RNA sample is reverse transcribed using MMLV reverse-transcriptase, 0.05 pg/µl oligo-(dT) primer (21mer), 1X first strand buffer, 0.03 units/µl RNase inhibitor, 500 µM dATP, 500 µM dGTP, 500 µM dTTP, 40 µM dCTP, 40 µM dCTP-Cy3 (BDS) or dCTP-Cy5 (Amersham Pharmacia Biotech). The reverse transcription reaction is performed in a 25 ml volume containing 200 ng poly(A)$^+$ RNA with GEMBRIGHT kits (Incyte). Specific control poly(A)$^+$ RNAs are synthesized by in vitro transcription from non-coding yeast genomic DNA. After incubation at 37°C for 2 hr, each reaction sample (one with Cy3 and another with Cy5 labeling) is treated with 2.5 ml of 0.5M sodium hydroxide and incubated for 20 minutes at 85°C to the stop the reaction and degrade the RNA. Samples are purified using two successive CHROMA SPIN 30 gel filtration spin columns (CLONTECH Laboratories, Inc. (CLONTECH), Palo Alto CA) and after combining, both reaction samples are ethanol precipitated using 1 ml of glycogen (1 mg/ml), 60 ml sodium acetate, and 300 ml of 100% ethanol. The sample is then dried to completion using a SpeedVAC (Savant Instruments Inc., Holbrook NY) and resuspended in 14 µl 5X SSC/0.2% SDS.

Microarray Preparation

Sequences of the present invention are used to generate array elements. Each array element is amplified from bacterial cells containing vectors with cloned cDNA inserts. PCR amplification uses primers complementary to the vector sequences flanking the cDNA insert. Array elements are amplified in thirty cycles of PCR from an initial quantity of 1-2 ng to a final quantity greater than 5 µg. Amplified array elements are then purified using SEPHACRYL-400 (Amersham Pharmacia Biotech).

Purified array elements are immobilized on polymer-coated glass slides. Glass microscope slides (Corning) are cleaned by ultrasound in 0.1% SDS and acetone, with extensive distilled water washes between and after treatments. Glass slides are etched in 4% hydrofluoric acid (VWR Scientific Products Corporation (VWR), West Chester PA), washed extensively in distilled water, and coated with 0.05% aminopropyl silane (Sigma) in 95% ethanol. Coated slides are cured in a 110°C oven.

Array elements are applied to the coated glass substrate using a procedure described in US Patent No. 5,807,522, incorporated herein by reference. 1 µl of the array element DNA, at an average concentration of 100 ng/µl, is loaded into the open capillary printing element by a high-speed robotic apparatus. The apparatus then deposits about 5 nl of array element sample per slide.

Microarrays are UV-crosslinked using a STRATALINKER UV-crosslinker (Stratagene). Microarrays are washed at room temperature once in 0.2% SDS and three times in distilled water. Non-specific binding sites are blocked by incubation of microarrays in 0.2% casein in phosphate buffered saline (PBS) (Tropix, Inc., Bedford MA) for 30 minutes at 60 °C followed by washes in 0.2% SDS and distilled water as before.

Hybridization

Hybridization reactions contain 9 µl of sample mixture consisting of 0.2 µg each of Cy3 and Cy5 labeled cDNA synthesis products in 5X SSC, 0.2% SDS hybridization buffer. The sample mixture is heated to 65 °C for 5 minutes and is aliquoted onto the microarray surface and covered with an 1.8 cm^2 coverslip. The arrays are transferred to a waterproof chamber having a cavity just slightly larger than a microscope slide. The chamber is kept at 100% humidity internally by the addition of 140 µl of 5X SSC in a corner of the chamber. The chamber containing the arrays is incubated for about 6.5 hours at 60 °C. The arrays are washed for 10 min at 45 °C in a first wash buffer (1X SSC, 0.1% SDS), three times for 10 minutes each at 45 °C in a second wash buffer (0.1X SSC), and dried.

Detection

Reporter-labeled hybridization complexes are detected with a microscope equipped with an Innova 70 mixed gas 10 W laser (Coherent, Inc., Santa Clara CA) capable of generating spectral lines at 488 nm for excitation of Cy3 and at 632 nm for excitation of Cy5. The excitation laser light is

focused on the array using a 20X microscope objective (Nikon, Inc., Melville NY). The slide containing the array is placed on a computer-controlled X-Y stage on the microscope and raster-scanned past the objective. The 1.8 cm x 1.8 cm array used in the present example is scanned with a resolution of 20 micrometers.

In two separate scans, a mixed gas multiline laser excites the two fluorophores sequentially. Emitted light is split, based on wavelength, into two photomultiplier tube detectors (PMT R1477, Hamamatsu Photonics Systems, Bridgewater NJ) corresponding to the two fluorophores. Appropriate filters positioned between the array and the photomultiplier tubes are used to filter the signals. The emission maxima of the fluorophores used are 565 nm for Cy3 and 650 nm for Cy5. Each array is typically scanned twice, one scan per fluorophore using the appropriate filters at the laser source, although the apparatus is capable of recording the spectra from both fluorophores simultaneously.

The sensitivity of the scans is typically calibrated using the signal intensity generated by a cDNA control species added to the sample mixture at a known concentration. A specific location on the array contains a complementary DNA sequence, allowing the intensity of the signal at that location to be correlated with a weight ratio of hybridizing species of 1:100,000. When two samples from different sources (e.g., representing test and control cells), each labeled with a different fluorophore, are hybridized to a single array for the purpose of identifying genes that are differentially expressed, the calibration is done by labeling samples of the calibrating cDNA with the two fluorophores and adding identical amounts of each to the hybridization mixture.

The output of the photomultiplier tube is digitized using a 12-bit RTI-835H analog-to-digital (A/D) conversion board (Analog Devices, Inc., Norwood MA) installed in an IBM-compatible PC computer. The digitized data are displayed as an image where the signal intensity is mapped using a linear 20-color transformation to a pseudocolor scale ranging from blue (low signal) to red (high signal). The data is also analyzed quantitatively. Where two different fluorophores are excited and measured simultaneously, the data are first corrected for optical crosstalk (due to overlapping emission spectra) between the fluorophores using each fluorophore's emission spectrum.

A grid is superimposed over the fluorescence signal image such that the signal from each spot is centered in each element of the grid. The fluorescence signal within each element is then integrated to obtain a numerical value corresponding to the average intensity of the signal. The software used for signal analysis is the GEMTOOLS gene expression analysis program (Incyte).

IX. Complementary Polynucleotides

Sequences complementary to the FLEXHT-encoding sequences, or any parts thereof, are used to detect, decrease, or inhibit expression of naturally occurring FLEXHT. Although use of oligonucleotides comprising from about 15 to 30 base pairs is described, essentially the same procedure is used with smaller or with larger sequence fragments. Appropriate oligonucleotides are designed

using OLIGO 4.06 software (National Biosciences) and the coding sequence of FLEXHT. To inhibit transcription, a complementary oligonucleotide is designed from the most unique 5' sequence and used to prevent promoter binding to the coding sequence. To inhibit translation, a complementary oligonucleotide is designed to prevent ribosomal binding to the FLEXHT-encoding transcript.

X. Expression of FLEXHT

Expression and purification of FLEXHT is achieved using bacterial or virus-based expression systems. For expression of FLEXHT in bacteria, cDNA is subcloned into an appropriate vector containing an antibiotic resistance gene and an inducible promoter that directs high levels of cDNA transcription. Examples of such promoters include, but are not limited to, the *trp-lac* (*tac*) hybrid promoter and the T5 or T7 bacteriophage promoter in conjunction with the *lac* operator regulatory element. Recombinant vectors are transformed into suitable bacterial hosts, e.g., BL21(DE3). Antibiotic resistant bacteria express FLEXHT upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of FLEXHT in eukaryotic cells is achieved by infecting insect or mammalian cell lines with recombinant Autographica californica nuclear polyhedrosis virus (AcMNPV), commonly known as baculovirus. The nonessential polyhedrin gene of baculovirus is replaced with cDNA encoding FLEXHT by either homologous recombination or bacterial-mediated transposition involving transfer plasmid intermediates. Viral infectivity is maintained and the strong polyhedrin promoter drives high levels of cDNA transcription. Recombinant baculovirus is used to infect Spodoptera frugiperda (Sf9) insect cells in most cases, or human hepatocytes, in some cases. Infection of the latter requires additional genetic modifications to baculovirus. (See Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945.)

In most expression systems, FLEXHT is synthesized as a fusion protein with, e.g., glutathione S-transferase (GST) or a peptide epitope tag, such as FLAG or 6-His, permitting rapid, single-step, affinity-based purification of recombinant fusion protein from crude cell lysates. GST, a 26-kilodalton enzyme from Schistosoma japonicum, enables the purification of fusion proteins on immobilized glutathione under conditions that maintain protein activity and antigenicity (Amersham Pharmacia Biotech). Following purification, the GST moiety can be proteolytically cleaved from FLEXHT at specifically engineered sites. FLAG, an 8-amino acid peptide, enables immunoaffinity purification using commercially available monoclonal and polyclonal anti-FLAG antibodies (Eastman Kodak). 6-His, a stretch of six consecutive histidine residues, enables purification on metal-chelate resins (QIAGEN). Methods for protein expression and purification are discussed in Ausubel (1995, supra, ch. 10 and 16). Purified FLEXHT obtained by these methods can be used directly in the assays in Examples XII and XIV.

XI. Functional Assays

FLEXHT function is assessed by expressing the sequences encoding FLEXHT at physiologically elevated levels in mammalian cell culture systems. cDNA is subcloned into a mammalian expression vector containing a strong promoter that drives high levels of cDNA expression. Vectors of choice include pCMV SPORT plasmid (Life Technologies) and pCR3.1 plasmid (Invitrogen), both of which contain the cytomegalovirus promoter. 5-10 μg of recombinant vector are transiently transfected into a human cell line, for example, an endothelial or hematopoietic cell line, using either liposome formulations or electroporation. 1-2 μg of an additional plasmid containing sequences encoding a marker protein are co-transfected. Expression of a marker protein provides a means to distinguish transfected cells from nontransfected cells and is a reliable predictor of cDNA expression from the recombinant vector. Marker proteins of choice include, e.g., Green Fluorescent Protein (GFP; Clontech), CD64, or a CD64-GFP fusion protein. Flow cytometry (FCM), an automated, laser optics-based technique, is used to identify transfected cells expressing GFP or CD64-GFP and to evaluate the apoptotic state of the cells and other cellular properties. FCM detects and quantifies the uptake of fluorescent molecules that diagnose events preceding or coincident with cell death. These events include changes in nuclear DNA content as measured by staining of DNA with propidium iodide; changes in cell size and granularity as measured by forward light scatter and 90 degree side light scatter; down-regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; alterations in expression of cell surface and intracellular proteins as measured by reactivity with specific antibodies; and alterations in plasma membrane composition as measured by the binding of fluorescein-conjugated Annexin V protein to the cell surface. Methods in flow cytometry are discussed in Ormerod, M.G. (1994) Flow Cytometry, Oxford, New York NY.

The influence of FLEXHT on gene expression can be assessed using highly purified populations of cells transfected with sequences encoding FLEXHT and either CD64 or CD64-GFP. CD64 and CD64-GFP are expressed on the surface of transfected cells and bind to conserved regions of human immunoglobulin G (IgG). Transfected cells are efficiently separated from nontransfected cells using magnetic beads coated with either human IgG or antibody against CD64 (DYNAL, Lake Success NY). mRNA can be purified from the cells using methods well known by those of skill in the art. Expression of mRNA encoding FLEXHT and other genes of interest can be analyzed by northern analysis or microarray techniques.

XII. Production of FLEXHT Specific Antibodies

FLEXHT substantially purified using polyacrylamide gel electrophoresis (PAGE; see, e.g., Harrington, M.G. (1990) Methods Enzymol. 182:488-495), or other purification techniques, is used to immunize rabbits and to produce antibodies using standard protocols.

Alternatively, the FLEXHT amino acid sequence is analyzed using LASERGENE software (DNASTAR) to determine regions of high immunogenicity, and a corresponding oligopeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Methods for selection of appropriate epitopes, such as those near the C-terminus or in hydrophilic regions are well described in the art. (See, e.g., Ausubel, 1995, supra, ch. 11.)

Typically, oligopeptides of about 15 residues in length are synthesized using an ABI 431A peptide synthesizer (Perkin-Elmer) using FMOC chemistry and coupled to KLH (Sigma-Aldrich, St. Louis MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) to increase immunogenicity. (See, e.g., Ausubel, 1995, supra.) Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for antipeptide and anti-FLEXHT activity by, for example, binding the peptide or FLEXHT to a substrate, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG.

XIII. Purification of Naturally Occurring FLEXHT Using Specific Antibodies

Naturally occurring or recombinant FLEXHT is substantially purified by immunoaffinity chromatography using antibodies specific for FLEXHT. An immunoaffinity column is constructed by covalently coupling anti-FLEXHT antibody to an activated chromatographic resin, such as CNBr-activated SEPHAROSE (Amersham Pharmacia Biotech). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

Media containing FLEXHT are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of FLEXHT (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/FLEXHT binding (e.g., a buffer of pH 2 to pH 3, or a high concentration of a chaotrope, such as urea or thiocyanate ion), and FLEXHT is collected.

XIV. Identification of Molecules Which Interact with FLEXHT

FLEXHT, or biologically active fragments thereof, are labeled with ^{125}I Bolton-Hunter reagent. (See, e.g., Bolton A.E. and W.M. Hunter (1973) Biochem. J. 133:529-539.) Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled FLEXHT, washed, and any wells with labeled FLEXHT complex are assayed. Data obtained using different concentrations of FLEXHT are used to calculate values for the number, affinity, and association of FLEXHT with the candidate molecules.

Alternatively, molecules interacting with FLEXHT are analyzed using the yeast two-hybrid system as described in Fields, S. and O. Song (1989, Nature 340:245-246), or using commercially available kits based on the two-hybrid system, such as the MATCHMAKER system (Clontech).

FLEXHT may also be used in the PATHCALLING process (CuraGen Corp., New Haven CT)

which employs the yeast two-hybrid system in a high-throughput manner to determine all interactions between the proteins encoded by two large libraries of genes (Nandabalan, K. et al. (2000) U.S. Patent No. 6,057,101).

Various modifications and variations of the described methods and systems of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with certain embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology or related fields are intended to be within the scope of the following claims.

Table 1

Protein SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
1	56	1841446	COLNNOT07	869045R6 (LUNGAST01), 1251087F6 (LUNGFET03), 1574990F6 (LNODNOT03), 1841446H1 (COLNNOT07), 1860233F6 (PROSNOT18), 1860233T6 (PROSNOT18), 4999024H1 (MYEPTXT02)
2	57	1850310	LUNGFET03	045378H1 (CORNNOT01), 178670H1 (PLACNOB01), 1731750F6 (BRSTTUT08), 1850310F6 (LUNGFET03), 1850310H1 (LUNGFET03), 1850310T6 (LUNGFET03), 1850465F6 (LUNGFET03), 2171482F6 (ENDCNOT03), 2771742H1 (COLANOT02), 3689884H1 (HEAANOT01), 5388293H1 (BRAINOT19)
3	58	1887020	BLADTUT07	997447R1 (KIDNTUT01), 1603324F6 (LUNGNOT15), 1887020F6 (BLADTUT07), 1887020H1 (BLADTUT07), 3149213H1 (ADRENON04), 3415447H1 (PTHYNOT04), 3472602H1 (LUNGNOT27), 5043722H2 (PLACFER01), 5300608H1 (MUSCNOT11)
4	59	1911421	CONNTUT01	276895H1 (TESTNOT03), 278868H1 (TESTNOT03), 1911421H1 (CONNTUT01), 1911421T6 (CONNTUT01), 2458650F6 (ENDANOT01)
5	60	1911910	CONNTUT01	1321511F1 (BLADNOT04), 1461469R1 (PANCNOT04), 1842277T6 (COLNNOT07), 1911910F6 (CONNTUT01), 1911910H1 (CONNTUT01)
6	61	1928920	BRSTNOT02	1928920H1 (BRSTNOT02), 637943F1 (BRSTNOT03), 1260702R1 (SYNORAT05), 1461588T1 (PANCNOT04), 1754346F6 (LIVRTUT01), 2906971F6 (THYMNOT05), 3368705F6 (CONNTUT04)
7	62	2170846	ENDCNOT03	2170846F6 (ENDCNOT03), 2170846H1 (ENDCNOT03), 2828087T6 (TLYMNOT03), SBLA00670F1
8	63	2176361	ENDCNOT03	660127X300D2 (BRAINOT03), 1212293R6 (BRSTTUT01), 1310486T1 (COLNFET02), 1841413H1 (COLNNOT07), 2176361H1 (ENDCNOT03), 2176361T6 (ENDCNOT03), 3212993T6 (BLADNOT08)
9	64	2212732	SINTFET03	285007F1 (EOSIHET02), 1002395R1 (BRSTNOT03), 2212732H1 (SINTFET03), 2361923R6 (LUNGFET05), 2708785T6 (PONSAZT01), 3605611H1 (LUNGNOT30)
10	65	2303457	BRSTNOT05	1573212F6 (LNODNOT03), 2303457H1 (BRSTNOT05), 2539058F6 (BONRTUT01), 3605342H1 (LUNGNOT30)
11	66	2317552	OVARNOT02	409183F1 (EOSIHET02), 863694R1 (BRAITUT03), 863694T1 (BRAITUT03), 913694H1 (STOMNOT02), 1319506T1 (BLADNOT04), 2317552H1 (OVARNOT02), 2604101H1 (LUNGTUT07)
12	67	2416366	HNT3AZT01	228599F1 (PANCNOT01), 1384425T1 (BRAITUT08), 1636237F6 (UTRSNOT06), 2416366H1 (HNT3AZT01), 2600055F6 (UTRSNOT10), 2868081F6 (KIDNNOT20), 2868081T6 (KIDNNOT20)
13	68	2472980	THP1NOT03	034079R6 (THP1NOB01), 034079T6 (THP1NOB01), 1595315F6 (BRAINOT14), 2472980H1 (THP1NOT03), 2596942T6 (OVARTUT02), 3852779T6 (BRAITUT12), 4754861H1 (BRAHNOT01)
14	69	2541640	BONRTUT01	621985F1 (PGANNOT01), 621985R6 (PGANNOT01), 2541640F6 (BONRTUT01), 2541640H1 (BONRTUT01), 5314848H1 (KIDETXS02)
15	70	2695204	UTRSNOT12	927986R1 (BRAINOT04), 1854107F6 (HNT3AZT01), 2695204H1 (UTRSNOT12), 3464157F6 (293TF2T01), 3596821H1 (FIBPNOT01)

Table 1 (cont.)

Protein SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
16	71	2805526	BLADTUT08	161563F1 (ADENINB01), 1275596F6 (TESTTUT02), 1508327F6 (LUNGNOT14), 1965858H1 (BRSTNOT04), 2207901F6 (SINTFET03), 2805526H1 (BLADTUT08), 3602953F6 (DRGTNOT01)
17	72	2850382	BRSTTUT13	1217874T1 (NEUTGMT01), 1267183F1 (BRAINOT09), 1515874F1 (PANCTUT01), 2234171F6 (PANCTUT02), 2850382H1 (BRSTTUT13), 3085014H1 (HEAONOT03)
18	73	2929276	TLYMNOT04	2929276H1 (TLYMNOT04) 136458R1 (SYNORAB01), 149704R1 (FIBRNGT02), 1231828X27 (BRAITUT01), 1236144F1 (LUNGFET03), 1351889F1 (LATRTUT02), 1381121F1 (BRAITUT08), 2509933F6 (CONUTUT01), 2819412H1 (BRSTNOT14), 3029720H1 (HEARFET02), 4897855H1 (OVARDIT01)
19	74	3033039	TLYMNOT05	3033039H1 (TLYMNOT05), 1235556F1 (LUNGFET03), 1251391F1 (LUNGFET03), 1261615H1 (SYNORAT05), 1413023F6 (BRAINOT12), 1425681T1 (BEPINON01), 1498538F1 (SINTBST01), 1562136F1 (SPLNNOT04), 1901480T6 (BLADTUT06), 2170485F6 (ENDCNOT03), 2454842F6 (ENDANOT01), 2915250H1 (THYMFET03), 4248873H1 (BRADDIR01), 4721266H1 (BRAIHCT02), 4939383H1 (BRAIFEN03), 5094946H1 (EPIMNON05), 5151444H1 (HEARFET03)
20	75	3213216	BLADNOT08	258371H1 (HNT2RAT01), 828313R1 (PROSNOT06), 914942R1 (BRSTNOT04), 917347R1 (BRSTNOT04), 1995075R6 (BRSTTUT03), 2854127H1 (CONNNOT02), 3213216H1 (BLADNOT08), 3323485F6 (PTHYNOT03), 3323485T6 (PTHYNOT03)
21	76	3220944	COLNNON03	015280F1 (HUVELPB01), 015280R1 (HUVELPB01), 1381368F6 (BRAITUT08), 1381368T6 (BRAITUT08), 3220944H1 (COLNNON03)
22	77	3224631	UTRSNON03	1321863F1 (BLADTUT04), 1517965H1 (BLADTUT04), 3224631H1 (UTRSNON03)
23	78	3242839	BRAINOT19	215253F1 (STOMNOT01), 2896446H1 (KIDNTUT14), 3242839H1 (BRAINOT19)
24	79	3274451	PROSBPT06	1879213F6 (LEUKNOT03), 2014292R6 (TESTNOT03), 3274451F6 (PROSBPT06), 3274451H1 (PROSBPT06), 3745237H1 (THYMNOT08)
25	80	3284256	HEAONOT05	3284256H1 (HEAONOT05), 4564779F6 (HELATXT01), 4564779T6 (HELATXT01)
26	81	3507004	CONCNOT01	116627R1 (KIDNNOT01), 1726869F6 (PROSNOT14), 3181596H1 (TLYJNOT01), 3507004H1 (CONCNOT01), 3703420H2 (PENCNOT07), SZAI03146F1, SZAI03025F1
27	82	3585823	293TF4T01	3585823H1 (293TF4T01), 825828R1 (PROSNOT06), 1871339F6 (SKINBIT01), 1568262F1 (UTRSNOT05), 1503644F6 (BRAITUT07)
28	83	3743822	THYMNOT08	2500776F6 (ADRETUT05), 3743822H1 (THYMNOT08), SBOA02104D1, SBOA02315D1, SBOA03346D1,
29	84	3808027	CONTTUT01	267411R1 (HNT2NOT01), 2579952H1 (KIDNTUT13), 3391551F6 (LUNGNOT28), 3483209H1 (KIDNNOT31), 3808027F6 (CONTTUT01), 3808027H1 (CONTTUT01), 5604051H1 (MONOTXN03)
30	85	4016220	BRAXNOT01	604824H1 (BRSTTUT01), 2290119H1 (BRAINON01), 4016220F6 (BRAXNOT01), 4016220H1 (BRAXNOT01), 4016220T6 (BRAXNOT01)
31	86	4093555	BSCNSZT01	404960F1 (TMLR3DT01), 413190F1 (BRSTNOT01), 1876856H1 (LEUKNOT03), 3623510H1 (ENDANOT03), 4093555H1 (BSCNSZT01)

WEST

Table 1 (cont.)

Protein SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
32	87	4829366	BRAVTXT03	1433606R1 (BEPINON01), 2457866T6 (ENDANOT01), 4829366H1 (BRAVTXT03)
33	88	4830720	BRAVTXT03	4830720H1 (BRAVTXT03), 114702F1 (TESTNOT01), 193078F1 (KIDNNOT02), 777503R6 (COLNNOT05)
34	89	5389730	BRAINOT19	1433103R1 (BEPINON01), 1553327F1 (BLADTUT04), 1555056T6 (BLADTUT04), 2794101F6 (COLNTUT16), 2936878H1 (THYMFET02), 5389730H1 (BRAINOT19)
35	90	5397088	LIVRTUT13	1626025F6 (COLNPOT01), 1626025T6 (COLNPOT01), 3586567H1 (293TF4T01), 3595435H1 (FIBPNOT01), 5397088H1 (LIVRTUT13), 5566591H1 (TLYMNOT08)
36	91	5425521	PROSTMT07	044182H1 (TBLYNOT01), 044182R6 (TBLYNOT01), 188753R6 (CARDNOT01), 533221R6 (BRAINOT03), 5425521H1 (PROSTMT07)
37	92	5495427	BRABDIR01	1213367R1 (BRSTTUT01), 2011728T6 (TESTNOT03), 5495427H1 (BRABDIR01)
38	93	003908	HMC1NOT01	003908H1 (HMC1NOT01), 1614119T6 (COLNTUT06), 1978612H1 (LUNGTUT03), 2743637F6 (BRSTTUT14), 2885045F6 (SINJNOT02), 3622505H1 (ENDANOT03), 3934887H1 (PROSTUT09)
39	94	085915	LIVRNOT01	085915H1 (LIVRNOT01), 085915R6 (LIVRNOT01), 085915T6 (LIVRNOT01), 2722468F6 (LUNGTUT10), 2809226F6 (BLADTUT08), 3211820F6 (BLADNOT08), 4028368H1 (BRAINOT23)
40	95	478615	MMLR2DT01	070322H1 (HUVESTB01), 160876F1 (ADENINB01), 478615H1 (MMLR2DT01), 508455H1 (TMLR3DT02), 835112R6 (PROSNOT07), 1393846F1 (THYRNOT03), 3853710H1 (BRAITUT12)
41	96	924880	BRAINOT04	491352R6 (HNT2AGT01), 491352T6 (HNT2AGT01), 924880H1 (BRAINOT04), 1970915F6 (UCMCL5T01), 2449392T6 (ENDANOT01), 3043413H1 (HEAANOT01)
42	97	955431	KIDNNOT05	157664F1 (THP1PLB02), 157664R1 (THP1PLB02), 955431H1 (KIDNNOT05), 2129525T6 (KIDNNOT05), 1962692R6 (BRSTNOT04), 3596009F6 (FIBPNOT01), SBIA10080D1
43	98	1275918	TESTTUT02	1275918H1 (TESTTUT02), 2858582H1 (SININOT03), SBFA00115F1, SBFA00940F1, SBFA00522F1
44	99	1290896	BRAINOT11	775874R1 (COLNNOT05), 898670R1 (BRSTTUT03), 1290896H1 (BRAINOT11), 2368210X16F1 (ADRENOT07), 3282892F6 (HEAONOT05), 5023045H1 (OVARNON03)
45	100	1342736	COLNTUT03	1342736H1 (COLNTUT03), 1525977F1 (UCMCL5T01), 2890930T6 (LUNGFET04)
46	101	1394209	THYRNOT03	1221991T1 (NEUTGMT01), 1394209H1 (THYRNOT03), SBFA01753F1
47	102	1394647	THYRNOT03	670296R6 (CRBLNOT01), 1394647F6 (THYRNOT03), 1394647H1 (THYRNOT03), 1399448F6 (BRAITUT08), 1522214F1 (BLADTUT04), 1943229R6 (HIPONOT01)
48	103	1436854	PANCNOT08	419380F1 (BRSTNOT01), 914745R1 (BRSTNOT04), 917341R1 (BRSTNOT04), 968442R1 (BRSTNOT05), 1000186R1 (BRSTNOT03), 1213294T1 (BRSTTUT01), 1325716T6 (LPARNOT02), 1436854F1 (PANCNOT08), 1436854H1 (PANCNOT08), 1820922F6 (GBLATUT01), 2236604H1 (PANCTUT02), 3035889H1 (TLYMNOT05), 3617358H1 (EPIPNOT01)

Table 1 (cont.)

Protein SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
49	104	1447955	PLACNOT02	036268R6 (HUVENOB01), 1447955F1 (PLACNOT02), 1447955H1 (PLACNOT02), 1447955R1 (PLACNOT02), 1870540F6 (SKINBIT01), 1870540T6 (SKINBIT01), 2556206F6 (THYMNOT03), 2792466H1 (COLNTUT16), 2884487H1 (SINJNOT02), 3236317H1 (COLNUCT03), 3433789H1 (PENCNOT05), 3943045H1 (SCORNOT04), 4175196H1 (SINTNOT21)
50	105	1454689	PENITUT01	069107H1 (HUVESTB01), 197052R6 (KIDNNOT02), 288467F1 (EOSIHET02), 1303613T6 (PLACNOT02), 1454689F7 (PENITUT01), 1454689H1 (PENITUT01), 1501227T6 (SINTBST01), 1714876F6 (UCMCNOT02), 2061637R6 (OVARNOT03)
51	106	1568009	UTRSNOT05	914061H1 (STOMNOT02), 1451614F6 (PENITUT01), 1468355H1 (PANCTUT02), 1568009H1 (UTRSNOT05), 1568009T6 (UTRSNOT05), 1726822T6 (PROSNOT14), 1816343F6 (PROSNOT20)
52	107	1677811	STOMFET01	1285190F6 (COLNNOT16), 1545363T1 (PROSTUT04), 1677811H1 (STOMFET01), 1855610T6 (HNT3AZT01), 2887165F6 (SINJNOT02)
53	108	1756951	PITUNOT03	198647R6 (KIDNNOT02), 198647T6 (KIDNNOT02), 1443882R1 (THYRNOT03), 1629702T6 (COLNPOT01), 1756951H1 (PITUNOT03), 1930992F6 (COLNTUT03), 3516457H1 (DENDNOT01), 3536962H1 (KIDNNOT25), 4984196H1 (HELATXT05)
54	109	1833547	BRAINON01	407043R6 (EOSIHET02), 1443002R1 (THYRNOT03), 1833547H1 (BRAINON01), 2040573H1 (HIPONON02), 2122559T6 (BRSTNOT07)
55	110	1833723	BRAINON01	1833723H1 (BRAINON01), 1833723R6 (BRAINON01), 1649301H1 (PROSTUT09), 1803139H1 (SINTNOT13), 1930604H1 (COLNTUT03)

Table 2

Polypeptide SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites
1	349	S53 T114 T147 S11 S15 S141 S149 T156 S225 T252 S299 T313 T342 S343	N250
2	169	T39 T88 T71 S119 T128	N126
3	316	T85 S194 T196 S80 T230	
4	220	T39 S10 T169 S77 S205 S208	N15
5	235	S13 T153 S128 S182 S134 T187 Y155	
6	487	S235 S3 S183 T149	L448
7	212	S59 T84 Y147	N57 N206
8	241	T167 S168 S48 T79 S129 S7 T36 S66 T86 T219	
9	375	T67 T111 S123 S132 T317 S8 S81 T173 S355 Y19	
10	429	S257 S40 T90 S117 S232 S234 S236 T242 S2 T242 T252 S256 S387	N88 N173 N283
11	329	S8 T10 S29 S47 S49 S57 S227 S260 S314 T86 S124 S277	N64 N114 N122 N250
12	476	S195 T246 T453 S459 S10 S20 S22 S34 S55 S57 S72 S77 S95 S97 S118 S122 S137 S168 S184 S211 T263 S330 T357 T392 S415 S443 S64 S83 S86 S106 S197 T308 T382 S431 T437 S438	N3 N376
13	366	S330 T119 T188 S60 T65 S140 S185 T298 S305 T326 T337 S344 S357	N39 N161 N202 N269 N273 N348 N352
14	152	T81 T148 S45	N28
15	233	T87 S18 T172 S186 S189 S203	N130
16	357	T118 S189 T62 T80 S115 T126 S247 S273 S328 S337	
17	251	T48 T224 S106 T114 S126 S161 T180 T202 S235	
18	105	S49 T69 T98 S15 S34	
19	876	S119 S179 T180 S226 T262 T386 S450 S494 S529 T542 S560 S585 T689 S716 T728 S770 S804 T35 T110 T123 S238 T248 T457 S574 S634 S693 T702 S733 S745 S806	N397 N644
20	254	S137 S143 S161 S170 T193 S203 T22 S203 T210 S213 S220 T242 Y124 Y250	
21	259	S10 S20 S35 S44 S183 S197 S235 S249 S253 S255 S28 T135 S171 S203 S204 S235	N199
22	65	T29 S10	
23	163	T29 S52 T61 S76 S102 A160	
24	199	S139 S153 S32 S49 S109 S120	N24
25	231	S15 T80 S175 Y38 Y146	N152 N189
26	412	S297 T29 S80 S145 S184 S201 T276 T299 T42 S70 T140 T292 T328 S352 S379 S404	N290 N402
27	272	T117 S160 T194 S198 S21 T58 S84 S86 S127 S140 T164 T202 T238 S255 S8 S160	
28	242	S66 S74 S232 T24 T102	

WEST

Table 2 (cont.)

Polypeptide SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites
29	285	S23 S27 T47 T137 S206 S241 S270 T5 S17 T95 T122 T175 T194	N192
30	89	T3 T16 S59 T53	
31	210	T60 S168 T192 T100 S118 S165	N160
32	271	S80 S16 T85 S99 S121 T134 T141 S154 S161 S166 T199 S49 S118	
33	389	T16 S63 S117 T130 T158 S250 T309 S188 T341 S375	N241 N266
34	228	S2 T11 S53 T63 T141 S154 T96 T189	
35	330	S42 T121 T160 T25 T54 S66 T157 T213 S218 T283	N158
36	307	T23 S197 T224 S258 T265 L304 T304	N208
37	163	T148 T98 T132	
38	202	S48 S50 S180 S14 T27	
39	209	S27 T28 T18 S30 S76 S157 S193 T10 S83 Y125 Y150	
40	314	T7 T50 S58 S157 S217 T9 T18 S71 S138 S151 S233	N56 N66 N95 N114
41	137	S42 S74 S133 S119	N28
42	245	S6	
43	179	S14 T76 T129 T65 S122	N105 G89
44	165	S38 S42 S29 T64 S73 T97 Y159	
45	177	S71 S107 S36	N122
46	215	S200 S201 S2 T44 T46 T51 S61 S65 T188 S29 S33 T118 T141 S148 T154 S170 S175 S179 T195	N91 N115 N193
47	133	S52 T35	
48	579	S5 S69 T77 S93 S127 S238 T381 S417 S419 S475 S552 S558 S426 S546 S569 Y446	N20 N136 N137 N173 N462
49	139	S14 S109 S56 Y65	
50	314	S2 S50 T138 T145 S212 S294 S176 S226 T227 S299	N222
51	355	S90 T99 T224 S345 S116 S275 S286 S324	
52	179	S133 S40 S143 S156 S26 S66 S115 S124 T171 Y139	
53	403	T382 S73 S84 S85 T113 T119 T168 T169 T272 S277 S311 S41 T69 S84 T127 T274 T300	N32 N52 N350
54	163	S5 T35 S54 S77 T31 Y107	N29
55	110	T8	

WEST

Table 3

Nucleotide SEQ ID NO:	Useful Fragment	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
56	596-640	Reproductive (0.235) Hematopoietic/Immune (0.148) Cardiovascular (0.136)	Cancer (0.432) Inflammation (0.284) Cell Proliferation (0.222)	PSPORT1
57	1380-1424	Reproductive (0.296) Nervous (0.167) Gastrointestinal (0.130)	Cancer (0.426) Inflammation (0.222) Cell Proliferation (0.204)	pINCY
58	273-317	Reproductive (0.200) Cardiovascular (0.175) Nervous (0.175)	Cancer (0.475) Cell Proliferation (0.175) Inflammation (0.175)	pINCY
59	111-155	Reproductive (0.215) Cardiovascular (0.139) Nervous (0.139)	Cancer (0.430) Inflammation (0.241) Cell Proliferation (0.228)	pINCY
60	217-261	Nervous (0.351) Gastrointestinal (0.135) Reproductive (0.135)	Cancer (0.405) Cell Proliferation (0.216) Inflammation (0.189)	pINCY
61	543-587	Reproductive (0.237) Gastrointestinal (0.175) Nervous (0.150)	Cancer (0.475) Inflammation (0.237) Cell Proliferation (0.225)	PSPORT1
62	138-182	Reproductive (0.250) Cardiovascular (0.200) Gastrointestinal (0.150)	Cancer (0.500) Cell Proliferation (0.250)	pINCY
63	1029-1073	Reproductive (0.325) Nervous (0.125) Urologic (0.125)	Cancer (0.450) Cell Proliferation (0.200) Inflammation (0.150)	pINCY
64	434-478	Nervous (0.250) Reproductive (0.222) Gastrointestinal (0.139)	Cancer (0.444) Inflammation (0.306) Cell Proliferation (0.194)	pINCY
65	327-372	Reproductive (0.333) Nervous (0.278) Cardiovascular (0.111)	Cancer (0.389) Inflammation (0.278) Cell Proliferation (0.111)	PSPORT1
66	992-1036	Nervous (0.235) Reproductive (0.235) Gastrointestinal (0.157)	Cancer (0.382) Inflammation (0.255) Cell Proliferation (0.235)	PSPORT1
67	325-369	Nervous (0.250) Reproductive (0.219) Hematopoietic/Immune (0.172)	Cancer (0.484) Inflammation (0.203) Cell Proliferation (0.141)	pINCY
68	336-380	Reproductive (0.219) Developmental (0.156) Hematopoietic/Immune (0.156)	Cancer (0.438) Inflammation (0.250) Cell Proliferation (0.219)	pINCY

Table 3 (cont.)

Nucleotide SEQ ID NO:	Useful Fragment	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
69	109-153	Reproductive (0.364) Gastrointestinal (0.273) Musculoskeletal (0.182)	Cancer (0.818) Cell Proliferation (0.091) Trauma (0.091)	pINCY
70	597-641	Nervous (0.211) Reproductive (0.211) Hematopoietic/Immune (0.132)	Cancer (0.395) Inflammation (0.263) Cell Proliferation (0.132)	pINCY
71	921-965	Reproductive (0.324) Nervous (0.135) Developmental (0.108)	Cancer (0.514) Inflammation (0.162) Cell Proliferation (0.135)	pINCY
72	381-425	Reproductive (0.269) Nervous (0.192) Hematopoietic/Immune (0.141)	Cancer (0.436) Inflammation (0.192) Cell Proliferation (0.167)	pINCY
73	109-153	Reproductive (0.281) Hematopoietic/Immune (0.132) Cardiovascular (0.114)	Cancer (0.412) Inflammation (0.333) Cell Proliferation (0.158)	pINCY
74	433-477	Nervous (0.206) Reproductive (0.206) Cardiovascular (0.150)	Cancer (0.477) Cell Proliferation (0.224) Inflammation (0.206)	pINCY
75	651-695	Reproductive (0.513) Cardiovascular (0.128) Endocrine (0.103)	Cancer (0.564) Inflammation (0.179) Cell Proliferation (0.128)	pINCY
76	97-141	Reproductive (0.314) Hematopoietic/Immune (0.200) Gastrointestinal (0.143)	Cancer (0.543) Cell Proliferation (0.286) Inflammation (0.229)	PSPORT1
77	272-316	Reproductive (0.333) Hematopoietic/Immune (0.267) Cardiovascular (0.133)	Cancer (0.600) Trauma (0.267) Cell Proliferation (0.200)	pINCY
78	20-64	Reproductive (0.232) Nervous (0.197) Gastrointestinal (0.169)	Cancer (0.423) Inflammation (0.246) Cell Proliferation (0.204)	pINCY
79	541-587	Hematopoietic/Immune (0.381) Gastrointestinal (0.190) Nervous (0.190)	Inflammation (0.524) Cancer (0.286) Cell Proliferation (0.095)	pINCY
80	334-378	Cardiovascular (0.200) Gastrointestinal (0.200) Reproductive (0.200)	Cancer (0.700) Cell Proliferation (0.300) Inflammation (0.100)	pINCY
81	1083-1127	Reproductive (0.263) Gastrointestinal (0.150) Hematopoietic/Immune (0.150)	Cancer (0.463) Inflammation (0.300) Cell Proliferation (0.287)	pINCY

Table 3 (cont.)

Nucleotide SEQ ID NO:	Useful Fragment	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
82	381-426	Reproductive (0.400) Nervous (0.240)	Cancer (0.520) Cell Proliferation (0.160) Inflammation (0.160)	pINCY
83	18-65	Reproductive (0.600) Endocrine (0.200) Hematopoietic/Immune (0.200)	Cancer (0.800) Inflammation (0.200) Cell Proliferation (0.000)	pINCY
84	980-1024	Reproductive (0.240) Hematopoietic/Immune (0.200) Nervous (0.200)	Cancer (0.520) Inflammation (0.280) Cell Proliferation (0.200)	pINCY
85	406-450	Nervous (0.600) Reproductive (0.400) Cardiovascular (0.000)	Cancer (0.600) Inflammation (0.200) Neurological (0.200)	pINCY
86	109-153	Reproductive (0.232) Nervous (0.224) Cardiovascular (0.120)	Cancer (0.472) Inflammation (0.224) Cell Proliferation (0.184)	pINCY
87	783-825	Reproductive (0.242) Nervous (0.167) Cardiovascular (0.130)	Cancer (0.423) Inflammation (0.223) Cell Proliferation (0.195)	pINCY
88	24-71	Reproductive (0.267) Gastrointestinal (0.154) Nervous (0.134)	Cancer (0.514) Cell Proliferation (0.202) Inflammation (0.174)	pINCY
89	41-85	Reproductive (0.264) Gastrointestinal (0.189) Nervous (0.151)	Cancer (0.585) Inflammation (0.264) Cell Proliferation (0.113)	pINCY
90	15-59	Reproductive (0.243) Nervous (0.162) Cardiovascular (0.135)	Cancer (0.459) Inflammation (0.216) Cell Proliferation (0.162)	pINCY
91	865-909	Nervous (0.231) Reproductive (0.231) Musculoskeletal (0.179)	Cancer (0.462) Inflammation (0.256) Cell Proliferation (0.154)	pINCY
92	45-89	Reproductive (0.238) Nervous (0.168) Gastrointestinal (0.152)	Cancer (0.480) Inflammation (0.188) Cell Proliferation (0.180)	pINCY
93	12-56	Gastrointestinal (0.189) Reproductive (0.189) Nervous (0.135)	Cancer (0.514) Cell Proliferation (0.297) Inflammation (0.189)	PBLUESCRIPT
94	389-433	Cardiovascular (0.192) Nervous (0.192) Reproductive (0.192)	Cancer (0.500) Inflammation (0.231) Cell Proliferation (0.077)	PBLUESCRIPT

Table 3 (cont.)

Nucleotide SEQ ID NO:	Useful Fragment	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
95	174-218	Reproductive (0.250) Nervous (0.219) Hematopoietic/Immune (0.188)	Cancer (0.406) Inflammation (0.312)	PSPORT1
96	219-263	Nervous (0.324) Reproductive (0.235) Cardiovascular (0.147)	Cancer (0.412) Inflammation (0.353) Cell Proliferation (0.118)	PSPORT1
97	218-262	Reproductive (0.266) Hematopoietic/Immune (0.183) Nervous (0.138)	Cancer (0.450) Inflammation (0.257) Cell Proliferation (0.239)	PSPORT1
98	327-371	Reproductive (0.429) Cardiovascular (0.143) Gastrointestinal (0.143)	Cancer (0.500) Inflammation (0.214) Cell Proliferation (0.143)	pINCY
99	236-280	Reproductive (0.219) Nervous (0.188) Gastrointestinal (0.135)	Cancer (0.510) Inflammation (0.198) Cell Proliferation (0.177)	pINCY
100	172-216	Developmental (0.250) Gastrointestinal (0.250) Dermatologic (0.125)	Cancer (0.500) Cell Proliferation (0.250) Inflammation (0.250)	pINCY
101	172-216	Hematopoietic/Immune (0.350) Reproductive (0.150)	Inflammation (0.350) Cancer (0.250) Cell Proliferation (0.200)	pINCY
102	346-390	Reproductive (0.278) Nervous (0.241) Cardiovascular (0.114)	Cancer (0.633) Cell Proliferation (0.304) Inflammation (0.127)	pINCY
103	720-764	Reproductive (0.267) Gastrointestinal (0.164) Hematopoietic/Immune (0.130)	Cancer (0.466) Inflammation (0.205) Cell Proliferation (0.199)	pINCY
104	154-198	Reproductive (0.316) Gastrointestinal (0.237) Nervous (0.105)	Cancer (0.447) Inflammation (0.316) Cell Proliferation (0.184)	pINCY
105	606-650	Hematopoietic/Immune (0.241) Cardiovascular (0.155) Nervous (0.155)	Cancer (0.448) Inflammation (0.328) Cell Proliferation (0.172)	pINCY
106	217-261	Reproductive (0.325) Gastrointestinal (0.175) Nervous (0.150)	Cancer (0.550) Inflammation (0.150) Trauma (0.150)	pINCY
107	497-541	Gastrointestinal (0.180) Reproductive (0.180) Nervous (0.140)	Cancer (0.540) Cell Proliferation (0.240) Trauma (0.120)	pINCY

Table 3 (cont.)

Nucleotide SEQ ID NO:	Useful Fragment	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
108	766-810	Reproductive (0.241) Hematopoietic/Immune (0.204) Cardiovascular (0.148)	Cancer (0.444) Inflammation (0.259) Cell Proliferation (0.222)	PSPORT1
109	109-153	Hematopoietic/Immune (0.189) Nervous (0.189) Reproductive (0.162)	Cancer (0.541) Inflammation (0.189) Cell Proliferation (0.162)	PSPORT1
110	138-182	Hematopoietic/Immune (0.211) Nervous (0.211) Reproductive (0.211)	Cancer (0.368) Inflammation (0.368) Cell Proliferation (0.184)	PSPORT1

Table 4

Nucleotide SEQ ID NO:	Library	Library Comment
56	COLNNOT07	Library was constructed using RNA isolated from colon tissue removed from a 60-year-old Caucasian male during a left hemicolectomy.
57	LUNGFET03	Library was constructed using RNA isolated from lung tissue removed from a Caucasian female fetus, who died at 20 weeks' gestation.
58	BLADTUT07	Library was constructed using RNA isolated from bladder tumor tissue removed from the anterior bladder wall of a 58-year-old Caucasian male during a radical cystectomy, radical prostatectomy, and gastrostomy. Pathology indicated a grade 3 transitional cell carcinoma in the left lateral bladder. Patient history included angina and emphysema. Family history included acute myocardial infarction, atherosclerotic coronary artery disease, and type II diabetes.
59	CONNTUT01	Library was constructed using RNA isolated from a soft tissue tumor removed from the clival area of the skull of a 30-year-old Caucasian female. Pathology indicated chondroid chordoma with neoplastic cells reactive for keratin.
60	CONNTUT01	Library was constructed using RNA isolated from a soft tissue tumor removed from the clival area of the skull of a 30-year-old Caucasian female. Pathology indicated chondroid chordoma with neoplastic cells reactive for keratin.
61	BRSTNOT02	Library was constructed using RNA isolated from diseased breast tissue removed from a 55-year-old Caucasian female during a unilateral extended simple mastectomy. Pathology indicated proliferative fibrocysytic changes characterized by apocrine metaplasia, sclerosing adenosis, cyst formation, and ductal hyperplasia without atypia. Pathology for the associated tumor tissue indicated an invasive grade 4 mammary adenocarcinoma. Patient history included atrial tachycardia and a benign neoplasm. Family history included cardiovascular and cerebrovascular disease.
62	ENDCNOT03	Library was constructed using RNA isolated from dermal microvascular endothelial cells removed from a neonatal Caucasian male.
63	ENDCNOT03	Library was constructed using RNA isolated from dermal microvascular endothelial cells removed from a neonatal Caucasian male.
64	SINTFET03	Library was constructed using RNA isolated from small intestine tissue removed from a Caucasian female fetus, who died at 20 weeks' gestation.
65	BRSTNOT05	Library was constructed using RNA isolated from breast tissue removed from a 58-year-old Caucasian female during a unilateral extended simple mastectomy. Pathology for the associated tumor tissue indicated multicentric invasive grade 4 lobular carcinoma. Patient history included skin cancer, rheumatic heart disease, osteoarthritis, and tuberculosis. Family history included cerebrovascular and cardiovascular disease, breast and prostate cancer, and type I diabetes.
66	OVARNOT02	Library was constructed using RNA isolated from ovarian tissue removed from a 59-year-old Caucasian female who died of a myocardial infarction. Patient history included cardiomyopathy, coronary artery disease, previous myocardial infarctions, hypercholesterolemia, hypotension, and arthritis.

Table 4 (cont.)

Nucleotide SEQ ID NO:	Library	Library Comment
67	HNT3AZT01	Library was constructed using RNA isolated from the hNT2 cell line (derived from a human teratocarcinoma that exhibited properties characteristic of a committed neuronal precursor). Cells were treated for three days with 0.35 micromolar 5-aza-2'-deoxycytidine (AZ).
68	THP1NOT03	Library was constructed using polyA RNA isolated from untreated THP-1 cells. THP-1 (ATCC TIB 202) is a human promonocyte line derived from the peripheral blood of a 1-year-old Caucasian male with acute monocytic leukemia (Int. J. Cancer (1980) 26:171).
69	BONRTUT01	Library was constructed using RNA isolated from rib tumor tissue removed from a 16-year-old Caucasian male during a rib osteotomy and a wedge resection of the lung. Pathology indicated a metastatic grade 3 (of 4) osteosarcoma, forming a mass involving the chest wall.
70	UTRSNOT12	Library was constructed using RNA isolated from uterine myometrial tissue removed from a 41-year-old Caucasian female during a vaginal hysterectomy with dilation and curettage. The endometrium was secretory and contained fragments of endometrial polyps. Benign endo- and ectocervical mucosa were identified in the endocervix. Pathology for the associated tumor tissue indicated uterine leiomyoma.
71	BLADTUT08	Library was constructed using RNA isolated from bladder tumor tissue removed from a 72-year-old Caucasian male during a radical cystectomy and prostatectomy. Pathology indicated an invasive grade 3 (of 3) transitional cell carcinoma in the right bladder base. Family history included myocardial infarction, cerebrovascular disease, and brain cancer.
72	BRSTTUT13	Library was constructed using RNA isolated from breast tumor tissue removed from the right breast of a 46-year-old Caucasian female during a unilateral extended simple mastectomy with breast reconstruction. Pathology indicated an invasive grade 3 adenocarcinoma, ductal type with apocrine features and greater than 50% intraductal component. Patient history included breast cancer.
73	TLYMNOT04	Library was constructed using 0.5 micrograms of polyA RNA isolated from activated Th1 cells. These cells were differentiated from umbilical cord CD4 T cells with IL-12 and B7-transfected COS cells, and then activated for six hours with anti-CD3 and anti-CD28 antibodies.
74	TLYMNOT05	Library was constructed using polyA RNA isolated from nonactivated Th2 cells. These cells were differentiated from umbilical cord CD4 T cells with IL-4 in the presence of anti-IL-12 antibodies and B7-transfected COS cells.
75	BLADNOT08	Library was constructed using RNA isolated from the bladder tissue of an 11-year-old black male, who died from a gunshot wound.

Table 4 (cont.)

Nucleotide SEQ ID NO:	Library	Library Comment
76	COLNNON03	Normalized colon library was constructed from 2.84 million independent clones from a colon tissue library. Starting RNA was made from colon tissue removed from a 60-year-old Caucasian male during a left hemicolectomy. The normalization and hybridization conditions were adapted from Soares et al. (Proc. Natl. Acad. Sci. USA (1994) 91:9228-9232), Swaroop et al. (Nucleic Acids Res. (1991) 19:1954) and Bonaldo et al. (Genome Res. (1996) 6:791-806), using a significantly longer (48 hour) reannealing hybridization period.
77	UTRSNON03	Normalized library was constructed from 6.4 million independent clones from a uterine tissue library. RNA was isolated from uterine myometrial tissue removed from a 41-year-old Caucasian female during a vaginal hysterectomy with dilation and curettage. The endometrium was secretory and contained fragments of endometrial polyps. Patient history included ventral hernia and a benign ovarian neoplasm. The normalization and hybridization conditions were adapted from Soares et al. (Proc. Natl. Acad. Sci. USA (1994) 91:9228).
78	BRAINOT19	Library was constructed using RNA isolated from diseased brain tissue removed from the left frontal lobe of a 27-year-old Caucasian male during a brain lobectomy. Pathology indicated a focal deep white matter lesion, characterized by marked gliosis, calcifications, and hemosiderin-laden macrophages, consistent with a remote perinatal injury. This tissue also showed mild to moderate generalized gliosis, predominantly subpial and subcortical, consistent with chronic seizure disorder. The left temporal lobe, including the mesial temporal structures, showed focal, marked pyramidal cell loss and gliosis in hippocampal sector CA1, consistent with mesial temporal sclerosis. GFAP was positive for astrocytes. Family history included brain cancer.
79	PROSBPT06	Library was constructed using RNA isolated from diseased prostate tissue remove from a 66-year-old Caucasian male during a radical prostatectomy and lymph node excision. Pathology indicated adenofibromatous hyperplasia. Patient history included hemiparesis, depressive disorder, sleep apnea, psoriasis, mitral valve prolapse, cerebrovascular disease, benign hypertension, and impotence. Family history included benign hypertension, cerebrovascular disease, and colon cancer.
80	HEAONOT05	Library was constructed using RNA isolated from aortic tissue removed from a 17-year-old Hispanic female, who died from a gunshot wound.
81	CONCNOT01	Library was constructed using RNA isolated from chest wall soft tissue removed from a 63-year-old Caucasian male during a chest wall lesion destruction. Pathology indicated surgical margins were free of tumor. Patient history included MEN (multiple endocrine neoplasia) syndrome type I, abnormal secretion of gastrin, calcium metabolism disease, chronic stomach ulcer with hemorrhage, lung cancer, and calculus of the kidney. Family history included prostate cancer, benign hypertension, stroke, atherosclerotic coronary artery disease, type II diabetes, hyperlipidemia, and cancer of an unspecified location.

Table 4 (cont.)

Nucleotide SEQ ID NO:	Library	Library Comment
82	293TF4T01	Library was constructed using RNA isolated from a transformed embryonal cell line (293-EBNA) derived from kidney epithelial tissue transfected with e2f1 and dp1. The cells were transformed with adenovirus 5 DNA.
83	THYMNOT08	Library was constructed using polyA RNA isolated from thymus tissue removed from a 4-month-old Caucasian male during a total thymectomy and open heart repair of atrioventricular canal defect using hypothermia. Pathology indicated a grossly normal thymus. The patient presented with a congenital heart anomaly, congestive heart failure, and Down syndrome. Patient history included abnormal thyroid function study and premature birth.
84	CONTTUT01	Library was constructed using RNA isolated from tumorous soft tissue of the left lateral thigh removed from a 34-year-old Caucasian female during a soft tissue excision. Pathology indicated metastatic grade 2 myxoid liposarcoma which formed multiple, lobulated, circumscribed masses situated in the subcutaneous adipose tissue. Patient history included a malignant soft tissue neoplasm of the leg. Family history included benign hypertension, acute leukemia, and type II diabetes.
85	BRAXNOT01	Library was constructed using RNA isolated from cerebellar tissue removed from a 70-year-old male.
86	BSCNSZT01	Library was constructed using RNA isolated from diseased caudate nucleus tissue removed from the brain of a 49-year-old male.
87	BRAVTXT03	Library was constructed using RNA isolated from treated astrocytes removed from the brain of a female fetus who died after 22 weeks'gestation. The cells were treated with tumor necrosis factor (TNF) alpha and interleukin 1 (IL-1), 10ng/ml each for 24 hours.
88	BRAVTXT03	Library was constructed using RNA isolated from treated astrocytes removed from the brain of a female fetus who died after 22 weeks'gestation. The cells were treated with tumor necrosis factor (TNF) alpha and interleukin 1 (IL-1), 10ng/ml each for 24 hours.
89	BRAINOT19	Library was constructed using RNA isolated from diseased brain tissue removed from the left frontal lobe of a 27-year-old Caucasian male during a brain lobectomy. Pathology indicated a focal deep white matter lesion, characterized by marked gliosis, calcifications, and hemosiderin-laden macrophages, consistent with a remote perinatal injury. This tissue also showed mild to moderate generalized gliosis, predominantly subpial and subcortical, consistent with chronic seizure disorder. The left temporal lobe, including the mesial temporal structures, showed focal, marked pyramidal cell loss and gliosis in hippocampal sector CA1, consistent with mesial temporal sclerosis. GFAP was positive for astrocytes. Family history included brain cancer.

Table 4 (cont.)

Nucleotide SEQ ID NO:	Library	Library Comment
90	LIVRTUT13	Library was constructed using RNA isolated from liver tumor tissue removed from a 62-year-old Caucasian female during partial hepatectomy and exploratory laparotomy. Pathology indicated metastatic intermediate grade neuroendocrine carcinoma, consistent with islet cell tumor, forming nodules ranging in size, in the lateral and medial left liver lobe. The pancreas showed fibrosis, chronic inflammation and fat necrosis consistent with pseudocyst. The gallbladder showed mild chronic cholecystitis. Patient history included malignant neoplasm of the pancreas tail, pulmonary embolism, hyperlipidemia, thrombophlebitis, joint pain in multiple joints, type II diabetes, benign hypertension, and cerebrovascular disease. Family history included pancreatic cancer, secondary liver cancer, benign hypertension, and hyperlipidemia.
91	PROSTMT07	Library was constructed using RNA isolated from diseased prostate tissue removed from a 73-year-old Caucasian male during radical prostatectomy, closed prostatic biopsy, and regional lymph node excision. Pathology indicated adenofibromatous hyperplasia. Pathology for the associated tumor tissue indicated adenocarcinoma, Gleason 3+3, involving the left side peripherally and anteriorly. The patient presented with elevated prostate-specific antigen. Family history included benign hypertension and cerebrovascular disease.
92	BRABDIR01	Library was constructed using RNA isolated from diseased cerebellum tissue removed from the brain of a 57-year-old Caucasian male, who died from a cerebrovascular accident. Patient history included Huntington's disease and emphysema.
93	HMC1NOT01	Library was constructed using RNA isolated from the HMC-1 human mast cell line derived from a 52-year-old female. Patient history included mast cell leukemia.
94	LIVRNOT01	Library was constructed at Stratagene, using RNA isolated from the liver tissue of a 49-year-old male.
95	MMLR2DT01	Library was constructed using RNA isolated from plastic adherent mononuclear cells isolated from buffy coat units obtained from unrelated male and female donors.
96	BRAINOT04	Library was constructed using RNA isolated from the brain tissue of a 44-year-old Caucasian male with a cerebral hemorrhage. The tissue, which contained coagulated blood, came from the choroid plexus of the right anterior temporal lobe. Family history included coronary artery disease and myocardial infarction.
97	KIDNNOT05	Library was constructed using RNA isolated from the kidney tissue of a 2-day-old Hispanic female, who died from cerebral anoxia. Family history included congenital heart disease.
98	TESTTUT02	Library was constructed using RNA isolated from testicular tumor removed from a 31-year-old Caucasian male during unilateral orchiectomy. Pathology indicated embryonal carcinoma.
99	BRAINOT11	Library was constructed using RNA isolated from brain tissue removed from the right temporal lobe of a 5-year-old Caucasian male during a hemispherectomy. Pathology indicated extensive polymicrogyria and mild to moderate gliosis (predominantly subpial and subcortical), consistent with chronic seizure disorder. Family history included a cervical neoplasm.

Table 4 (cont.)

Nucleotide SEQ ID NO:	Library	Library Comment
100	COLNTUT03	Library was constructed using RNA isolated from colon tumor tissue obtained from the sigmoid colon of a 62-year-old Caucasian male during a sigmoidectomy and permanent colostomy. Pathology indicated invasive grade 2 adenocarcinoma. Family history included benign hypertension, atherosclerotic coronary artery disease, hyperlipidemia, breast cancer, and prostate cancer.
101	THYRNOT03	Library was constructed using RNA isolated from thyroid tissue removed from the left thyroid of a 28-year-old Caucasian female during a complete thyroidectomy. Pathology indicated a small nodule of adenomatous hyperplasia present in the left thyroid.
102	THYRNOT03	Library was constructed using RNA isolated from thyroid tissue removed from the left thyroid of a 28-year-old Caucasian female during a complete thyroidectomy. Pathology indicated a small nodule of adenomatous hyperplasia present in the left thyroid.
103	PANCNOT08	Library was constructed using RNA isolated from pancreatic tissue removed from a 65-year-old Caucasian female during radical subtotal pancreatectomy. Pathology for the associated tumor tissue indicated an invasive grade 2 adenocarcinoma. Family history included cardiovascular disease, type II diabetes, and stomach cancer.
104	PLACNOT02	Library was constructed using RNA isolated from the placental tissue of a Hispanic female fetus, who was prematurely delivered at 21 weeks' gestation. Serologies of the mother's blood were positive for CMV (cytomegalovirus).
105	PENITUT01	Library was constructed using RNA isolated from tumor tissue removed from the penis of a 64-year-old Caucasian male during penile amputation. Pathology indicated a fungating invasive grade 4 squamous cell carcinoma involving the inner wall of the foreskin and extending onto the glans penis. Family history included malignant pharyngeal neoplasm, chronic lymphocytic leukemia, and chronic liver disease.
106	UTRSNOT05	Library was constructed using RNA isolated from the uterine tissue of a 45-year-old Caucasian female during a total abdominal hysterectomy and total colectomy. Pathology for the associated tumor tissue indicated multiple leiomyomas of the myometrium and a grade 2 colonic adenocarcinoma of the cecum. Patient history included multiple sclerosis and mitral valve disorder. Family history included type I diabetes, cerebrovascular disease, atherosclerotic coronary artery disease, malignant skin neoplasm, hypertension, and malignant neoplasm of the colon.
107	STOMFET01	Library was constructed using RNA isolated from the stomach tissue of a Caucasian female fetus, who died at 20 weeks' gestation.
108	PITUNOT03	Library was constructed using polyA RNA isolated from pituitary tissue of a 46-year-old Caucasian male, who died from colon cancer. Serologies were negative. Patient history included arthritis and peptic ulcer disease.
109	BRAINON01	Library was constructed and normalized from 4.88 million independent clones from a brain tissue library. RNA was made from brain tissue removed from a 26-year-old Caucasian male during cranioplasty and excision of a cerebral meningeal lesion. Pathology for the associated tumor tissue indicated a grade 4 oligoastrocytoma in the right fronto-parietal part of the brain.

Table 4 (cont.)

Nucleotide SEQ ID NO:	Library	Library Comment
110	BRAINON01	Library was constructed and normalized from 4.88 million independent clones from a brain tissue library. RNA was made from brain tissue removed from a 26-year-old Caucasian male during cranioplasty and excision of a cerebral meningeal lesion. Pathology for the associated tumor tissue indicated a grade 4 oligoastrocytoma in the right fronto-parietal part of the brain.

Table 5

Polypeptide SEQ ID NO:	Identity	Homologous Sequences, Signature Sequences, Domains and Motifs	Analytical Methods and Databases
1	U3 small nucleolar ribonucleoprotein (Maxwell, E.S. and M.J. Fournier (1995) Annu. Rev. Biochem. 64:897-933)	Putative U3 small nucleolar ribonucleoprotein protein [Arabadopsis thaliana] (g6358792)	BLAST-GenBank
3	Sarcosine oxidase (Reuber, B.E. et al. (1997) J. Biol. Chem. 272:6766-6776)	Sarcosine oxidase subunit beta [Pyrococcus abyssi] (g5458207) Sarcosine oxidase domain: PD001376: K146-G307	BLAST-GenBank BLAST-PRODOM
4	Pop4 (RNase MRP subunit) (van Eenennaam, H. et al. (1999) Nucleic Acids Res. 27:2465-2472)	Pop4 [Homo sapiens] (g4493697) Pop4 domain: PD125728: R95-L220	BLAST-GenBank BLAST-PRODOM
6	DLTB (transport of activated D-alanine across membranes) (Neuhaus, F.C. et al. (1996) Microb. Drug Resist. 2:77-84)	Transmembrane domains: L102-M122, I454-I472 DLTB domain: PD007034: G221-H459	HMMER BLAST-PRODOM
7	Peroxisomal membrane protein (Reguenga, C. et al. (1999) Biochim. Biophys. Acta 1445:337-341)	Peroxisomal membrane protein PMP24 [Homo sapiens] (g5441952)	BLAST-GenBank
11	Secreted protein	Signal peptide: M1-G51	SPScan
13	Caffeine-induced death protein (protects cells from caffeine, which can override the replication checkpoint of the cell cycle) (Wang, S.W. et al. (1999) J. Cell Sci. 112:927-937)	Caffeine-induced death protein 1 [Schizosaccharomyces pombe] (g4324457)	BLAST-GenBank
14	Secreted protein	Signal peptide: M1-G32	SPScan
16	RNA-binding protein	KH domain (RNA-binding domain): T62-G107	HMMER-Pfam
17	Secreted protein	Signal peptide: M1-A31	SPScan
19	Thymopoietin (Weber, P.J. et al. (1999) Biol. Chem. 380:653-660)	ZTMPO-1 [Homo sapiens] (Sheppard, P.O. et al. WO99/54468-A1)	BLAST-GeneSeq
26	Adenocarcinoma antigen	ART-4 (adenocarcinoma antigen recognized by T-lymphocytes) [Homo sapiens] (g6467119)	BLAST-GenBank
30	Apoptosis signaling protein (Hofmann, K. et al. (1997) Trends Biochem. Sci. 22:155-156)	ASC (caspase recruitment domain protein) [Homo sapiens] (g6482372)	BLAST-GenBank

Table 5 (cont.)

Polypeptide SEQ ID NO:	Identity	Homologous Sequences, Signature Sequences, Domains and Motifs	Analytical Methods and Databases
32	Prostate cancer-associated protein	Prostate cancer-associated protein [Homo sapiens] (DE19811194-A1)	BLAST-GeneSeq
33	Secreted transmembrane protein	Signal peptide: M1-Y48 Transmembrane protein domains: PD024190: L7-W171 PD013994: D177-H311, I258-I379	SPScan BLAST-PRODOM
36	LIP5	LIP5 (LYST-interacting protein; LYST = lysosomal trafficking regulator, responsible for Chediak-Higashi syndrome) [Homo sapiens] (Nandabalan, K. and S Kingsmore WO99/51741-A2)	BLAST-GeneSeq
37	Breast tumor-associated protein	Breast tumor-associated protein [Homo sapiens] (DE19813839-A1)	BLAST-GeneSeq
40	Secreted protein	Signal peptide: M1-G54	SPScan
43	Secreted protein	Signal peptide: M1-S56	SPScan
45	Phosphatidylinositol transfer protein (Bankaitis, V.A. et al. (1990) Nature 347:561-562)	SEC14-like protein [Homo sapiens] (Q92503) MSF1 (SEC14-like protein) domain: PD007507: V10-E169	BLAST-SwissProt BLAST-PRODOM
46	Secreted protein	Signal peptide: M1-A19	SPScan
48	Renal-cell carcinoma antigen (Scanlan, M.J. et al. (1999) Int. J. Cancer 83:456-464)	NY-REN-2 antigen [Homo sapiens] (g5360085)	BLAST-GenBank
51	F-box protein (Cenciarelli, C. et al. (1999) Curr. Biol. 9:1177-1179)	F-box protein Fbx25 [Homo sapiens] (g6164753)	BLAST-GenBank
53	Renal-cell carcinoma antigen (Scanlan, M.J. et al. (1999) Int. J. Cancer 83:456-464)	NY-REN-49 antigen [Homo sapiens] (g5360117)	BLAST-GenBank
55	DNA repair cofactor (Bravo, R. et al. (1987) Nature 326:515-517)	Proliferating cell nuclear antigen signature: E63-P108	ProfileScan

Table 6

Program	Description	Reference	Parameter Threshold
ABI FACTURA	A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA.	
ABI/PARACEL FDF	A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA.	Mismatch <50%
ABI AutoAssembler	A program that assembles nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA.	
BLAST	A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastp, blastn, blastx, tblastn, and tblastx.	Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. et al. (1997) Nucleic Acids Res. 25: 3389-3402.	*ESTs:* Probability value= 1.0E-8 or less *Full Length sequences:* Probability value= 1.0E-10 or less
FASTA	A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type. FASTA comprises as least five functions: fasta, tfasta, fastx, tfastx, and ssearch.	Pearson, W.R. and D.J. Lipman (1988) Proc. Natl. Acad Sci. 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183: 63-98; and Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489.	*ESTs:* fasta E value=1.06E-6 *Assembled ESTs:* fasta Identity= 95% or greater and Match length=200 bases or greater; fastx E value=1.0E-8 or less *Full Length sequences:* fastx score=100 or greater
BLIMPS	A BLocks IMProved Searcher that matches a sequence against those in BLOCKS, PRINTS, DOMO, PRODOM, and PFAM databases to search for gene families, sequence homology, and structural fingerprint regions.	Henikoff, S and J.G. Henikoff, Nucl. Acid Res., 19:6565-72, 1991. J.G. Henikoff and S. Henikoff (1996) Methods Enzymol. 266:88-105; and Attwood, T.K. et al. (1997) J. Chem. Inf. Comput. Sci. 37: 417-424.	Score=1000 or greater; Ratio of Score/Strength = 0.75 or larger; and, if applicable, Probability value= 1.0E-3 or less
HMMER	An algorithm for searching a query sequence against hidden Markov model (HMM)-based databases of protein family consensus sequences, such as PFAM.	Krogh, A. et al. (1994) J. Mol. Biol., 235:1501-1531; Sonnhammer, E.L.L. et al. (1988) Nucleic Acids Res. 26:320-322.	Score=10-50 bits for PFAM hits, depending on individual protein families

Table 6 (cont.)

Program	Description	Reference	Parameter Threshold
ProfileScan	An algorithm that searches for structural and sequence motifs in protein sequences that match sequence patterns defined in Prosite.	Gribskov, M. et al. (1988) CABIOS 4:61-66; Gribskov, et al. (1989) Methods Enzymol. 183:146-159; Bairoch, A. et al. (1997) Nucleic Acids Res. 25: 217-221.	Normalized quality score≥GCG-specified "HIGH" value for that particular Prosite motif. Generally, score=1.4-2.1.
Phred	A base-calling algorithm that examines automated sequencer traces with high sensitivity and probability.	Ewing, B. et al. (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	
Phrap	A Phils Revised Assembly Program including SWAT and CrossMatch, programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence homology and assembling DNA sequences.	Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M. S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.	Score= 120 or greater; Match length= 56 or greater
Consed	A graphical tool for viewing and editing Phrap assemblies	Gordon, D. et al. (1998) Genome Res. 8:195-202.	
SPScan	A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Nielson, H. et al. (1997) Protein Engineering 10:1-6; Claverie, J.M. and S. Audic (1997) CABIOS 12: 431-439.	Score=3.5 or greater
Motifs	A program that searches amino acid sequences for patterns that matched those defined in Prosite.	Bairoch et al. supra; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	

What is claimed is:

1. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of:

a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-55,

b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-55,

c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-55, and

d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-55.

2. An isolated polypeptide of claim 1 selected from the group consisting of SEQ ID NO:1-55.

3. An isolated polynucleotide encoding a polypeptide of claim 1.

4. An isolated polynucleotide encoding a polypeptide of claim 2.

5. An isolated polynucleotide of claim 4 selected from the group consisting of SEQ ID NO:56-110.

6. A recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide of claim 3.

7. A cell transformed with a recombinant polynucleotide of claim 6.

8. A transgenic organism comprising a recombinant polynucleotide of claim 6.

9. A method for producing a polypeptide of claim 1, the method comprising:

a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide, and said recombinant polynucleotide comprises a promoter sequence operably linked to a polynucleotide encoding the polypeptide of claim 1, and

b) recovering the polypeptide so expressed.

10. An isolated antibody which specifically binds to a polypeptide of claim 1.

11. An isolated polynucleotide comprising a polynucleotide sequence selected from the group consisting of:

a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:56-110,

b) a naturally occurring polynucleotide sequence having at least 90% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:56-110,

c) a polynucleotide sequence complementary to a),

d) a polynucleotide sequence complementary to b), and

e) an RNA equivalent of a)-d).

12. An isolated polynucleotide comprising at least 60 contiguous nucleotides of a polynucleotide of claim 11.

13. A method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 11, the method comprising:

a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and

b) detecting the presence or absence of said hybridization complex, and, optionally, if present, the amount thereof.

14. A method of claim 13, wherein the probe comprises at least 60 contiguous nucleotides.

15. A method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 11, the method comprising:

a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and

b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.

16. A pharmaceutical composition comprising an effective amount of a polypeptide of claim 1 and a pharmaceutically acceptable excipient.

17. A pharmaceutical composition of claim 16, wherein the polypeptide comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-55.

18. A method for treating a disease or condition associated with decreased expression of functional FLEXHT, comprising administering to a patient in need of such treatment the pharmaceutical composition of claim 16.

19. A method for screening a compound for effectiveness as an agonist of a polypeptide of claim 1, the method comprising:

a) exposing a sample comprising a polypeptide of claim 1 to a compound, and

b) detecting agonist activity in the sample.

20. A pharmaceutical composition comprising an agonist compound identified by a method of claim 19 and a pharmaceutically acceptable excipient.

21. A method for treating a disease or condition associated with decreased expression of functional FLEXHT, comprising administering to a patient in need of such treatment a pharmaceutical composition of claim 20.

22. A method for screening a compound for effectiveness as an antagonist of a polypeptide of claim 1, the method comprising:

a) exposing a sample comprising a polypeptide of claim 1 to a compound, and

b) detecting antagonist activity in the sample.

23. A pharmaceutical composition comprising an antagonist compound identified by a method of claim 22 and a pharmaceutically acceptable excipient.

24. A method for treating a disease or condition associated with overexpression of functional FLEXHT, comprising administering to a patient in need of such treatment a pharmaceutical composition of claim 23.

25. A method of screening for a compound that specifically binds to the polypeptide of claim 1, said method comprising the steps of:

a) combining the polypeptide of claim 1 with at least one test compound under suitable conditions, and

b) detecting binding of the polypeptide of claim 1 to the test compound, thereby identifying a compound that specifically binds to the polypeptide of claim 1.

26. A method of screening for a compound that modulates the activity of the polypeptide of claim 1, said method comprising:

a) combining the polypeptide of claim 1 with at least one test compound under conditions permissive for the activity of the polypeptide of claim 1,

b) assessing the activity of the polypeptide of claim 1 in the presence of the test compound, and

c) comparing the activity of the polypeptide of claim 1 in the presence of the test compound with the activity of the polypeptide of claim 1 in the absence of the test compound, wherein a change in the activity of the polypeptide of claim 1 in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide of claim 1.

27. A method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a sequence of claim 5, the method comprising:

a) exposing a sample comprising the target polynucleotide to a compound, and

b) detecting altered expression of the target polynucleotide.

28. An isolated polynucleotide comprising at least 20 contiguous nucleotides of a polynucleotide having a polynucleotide sequence selected from the group consisting of:

a) a polynucleotide sequence which is complementary to a polynucleotide sequence selected from the group consisting of SEQ ID NO:56-110, and

b) an RNA equivalent of a).

29. A composition for the detection of altered expression of human FLEXHT polynucleotides comprising at least one of the polynucleotides of claim 28 and a detectable label.

30. A microarray wherein at least one element of the microarray is a polynucleotide of claim 28.

31. A method for generating a transcript image of a sample which contains polynucleotides, the method comprising the steps of:

a) labeling the polynucleotides of the sample,

b) contacting the elements of the microarray of claim 30 with the labeled polynucleotides of the sample under conditions suitable for the formation of a hybridization complex, and

c) quantifying the expression of the polynucleotides in the sample.

SEQUENCE LISING

<110> INCYTE GENOMICS, INC.
YUE, Henry
TANG, Y. Tom
LAL, Preeti
REDDY, Roopa
BATRA, Sajeev
BAUGHN, Mariah R.
YANG, Junming
AZIMZAI, Yalda
LU, Dyung Aina M.
AU-YOUNG, Janice
SHIH, Leo L.

<120> FULL-LENGTH MOLECULES EXPRESSED IN HUMAN TISSUES

<130> PF-0695 PCT

<140> To Be Assigned
<141> Herewith

<150> 09/311,894; 09/311,940; 09/311,937
<151> 1999-05-14; 1999-05-14; 1999-05-14

<160> 110

<170> PERL Program

<210> 1
<211> 349
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1841446CD1

<400> 1
```
Met Ala Lys Ala Gly Asp Lys Ser Ser Ser Ser Gly Lys Lys Ser
 1               5                  10                  15
Leu Lys Arg Lys Ala Ala Ala Glu Glu Leu Gln Glu Ala Ala Gly
                20                  25                  30
Ala Gly Asp Gly Ala Thr Glu Asn Gly Val Gln Pro Pro Lys Ala
                35                  40                  45
Ala Ala Phe Pro Pro Gly Phe Ser Ile Ser Glu Ile Lys Asn Lys
                50                  55                  60
Gln Arg Arg His Leu Met Phe Thr Arg Trp Lys Gln Gln Gln Arg
                65                  70                  75
Lys Glu Lys Leu Ala Ala Lys Lys Lys Leu Lys Lys Glu Arg Glu
                80                  85                  90
Ala Leu Gly Asp Lys Ala Pro Pro Lys Pro Val Pro Lys Thr Ile
                95                 100                 105
Asp Asn Gln Arg Val Tyr Asp Glu Thr Thr Val Asp Pro Asn Asp
               110                 115                 120
Glu Glu Val Ala Tyr Asp Glu Ala Thr Asp Glu Phe Ala Ser Tyr
               125                 130                 135
Phe Asn Lys Gln Thr Ser Pro Lys Ile Leu Ile Thr Thr Ser Asp
               140                 145                 150
Arg Pro His Gly Arg Thr Val Arg Leu Cys Glu Gln Leu Ser Thr
               155                 160                 165
Val Ile Pro Asn Ser His Val Tyr Tyr Arg Arg Gly Leu Ala Leu
               170                 175                 180
Lys Lys Ile Ile Pro Gln Cys Ile Ala Arg Asp Phe Thr Asp Leu
               185                 190                 195
Ile Val Ile Asn Glu Asp Arg Lys Thr Pro Asn Gly Leu Ile Leu
```

200 205 210
Ser His Leu Pro Asn Gly Pro Thr Ala His Phe Lys Met Ser Ser
215 220 225
Val Arg Leu Arg Lys Glu Ile Lys Arg Arg Gly Lys Asp Pro Thr
230 235 240
Glu His Ile Pro Glu Ile Ile Leu Asn Asn Phe Thr Thr Arg Leu
245 250 255
Gly His Ser Ile Gly Arg Met Phe Ala Ser Leu Phe Pro His Asn
260 265 270
Pro Gln Phe Ile Gly Arg Gln Val Ala Thr Phe His Asn Gln Arg
275 280 285
Asp Tyr Ile Phe Phe Arg Phe His Arg Tyr Ile Phe Arg Ser Glu
290 295 300
Lys Lys Val Gly Ile Gln Glu Leu Gly Pro Arg Phe Thr Leu Lys
305 310 315
Leu Arg Ser Leu Gln Lys Gly Thr Phe Asp Ser Lys Tyr Gly Glu
320 325 330
Tyr Glu Trp Val His Lys Pro Arg Glu Met Asp Thr Ser Arg Arg
335 340 345
Lys Phe His Leu

<210> 2
<211> 169
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1850310CD1

<400> 2
Met Gln Cys Leu Leu Pro Tyr Gln Ser Lys Glu Pro Ser Cys Leu
1 5 10 15
Pro Pro Leu Pro Leu Asn Leu Pro Leu Pro Pro Cys Leu Cys Pro
20 25 30
Leu Leu Gln Leu Asn Ala Ala Met Thr Arg Lys Glu Lys Thr Lys
35 40 45
Glu Gly Gln Arg Ala Ala Gln Phe Ser Ala Gly Ala Asp Ala Gly
50 55 60
Ser Gly Gly Gly Leu Ser Arg Gln Lys Asp Thr Lys Arg Pro Met
65 70 75
Leu Leu Val Ile His Asp Val Val Leu Glu Leu Leu Thr Ser Ser
80 85 90
Asp Cys His Ala Asn Pro Arg Lys Tyr Pro Thr Cys Gln Lys Ser
95 100 105
Glu Val Leu Gly Val Ser Ile Tyr Val Ser Ile Cys Pro Ser Thr
110 115 120
Arg Pro Arg Asp Lys Asn Lys Thr Lys Lys Arg Cys Gln Val Leu
125 130 135
Glu Ala Val Leu Val Ser Lys Pro Ser Gly Ser Cys His Gln Gly
140 145 150
Ser Phe Glu Ile Val Pro His Val Lys Gly Asn Leu Ala Phe Thr
155 160 165
Ser Ser Asn His

<210> 3
<211> 316
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1887020CD1

<400> 3

Met Glu Ser Asn Val Lys Val Gln Arg Gln Glu Gly Ala Lys Val
1 5 10 15
Ser Leu Met Ser Pro Asp Gln Leu Arg Asn Lys Phe Pro Trp Ile
20 25 30
Asn Thr Glu Gly Val Ala Leu Ala Ser Tyr Gly Met Glu Asp Glu
35 40 45
Gly Trp Phe Asp Pro Trp Cys Leu Leu Gln Gly Leu Arg Arg Lys
50 55 60
Val Gln Ser Leu Gly Val Leu Phe Cys Gln Gly Glu Val Thr Arg
65 70 75
Phe Val Ser Ser Ser Gln Arg Met Leu Thr Thr Asp Asp Lys Ala
80 85 90
Val Val Leu Lys Arg Ile His Glu Val His Val Lys Met Asp Arg
95 100 105
Ser Leu Glu Tyr Gln Pro Val Glu Cys Ala Ile Val Ile Asn Ala
110 115 120
Ala Gly Ala Trp Ser Ala Gln Ile Ala Ala Leu Ala Gly Val Gly
125 130 135
Glu Gly Pro Pro Gly Thr Leu Gln Gly Thr Lys Leu Pro Val Glu
140 145 150
Pro Arg Lys Arg Tyr Val Tyr Val Trp His Cys Pro Gln Gly Pro
155 160 165
Gly Leu Glu Thr Pro Leu Val Ala Asp Thr Ser Gly Ala Tyr Phe
170 175 180
Arg Arg Glu Gly Leu Gly Ser Asn Tyr Leu Gly Gly Arg Ser Pro
185 190 195
Thr Glu Gln Glu Glu Pro Asp Pro Ala Asn Leu Glu Val Asp His
200 205 210
Asp Phe Phe Gln Asp Lys Val Trp Pro His Leu Ala Leu Arg Val
215 220 225
Pro Ala Phe Glu Thr Leu Lys Val Gln Ser Ala Trp Ala Gly Tyr
230 235 240
Tyr Asp Tyr Asn Thr Phe Asp Gln Asn Gly Val Val Gly Pro His
245 250 255
Pro Leu Val Val Asn Met Tyr Phe Ala Thr Gly Phe Ser Gly His
260 265 270
Gly Leu Gln Gln Ala Pro Gly Ile Gly Arg Ala Val Ala Glu Met
275 280 285
Val Leu Lys Gly Arg Phe Gln Thr Ile Asp Leu Ser Pro Phe Leu
290 295 300
Phe Thr Arg Phe Tyr Leu Gly Glu Lys Ile Gln Glu Asn Asn Ile
305 310 315
Ile

<210> 4
<211> 220
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1911421CD1

<400> 4
Met Lys Ser Val Ile Tyr His Ala Leu Ser Gln Lys Glu Ala Asn
1 5 10 15
Asp Ser Asp Val Gln Pro Ser Gly Ala Gln Arg Ala Glu Ala Phe
20 25 30
Val Arg Ala Phe Leu Lys Arg Ser Thr Pro Arg Met Ser Pro Gln
35 40 45
Ala Arg Glu Asp Gln Leu Gln Arg Lys Ala Val Val Leu Glu Tyr
50 55 60
Phe Thr Arg His Lys Arg Lys Glu Lys Lys Lys Lys Ala Lys Gly
65 70 75
Leu Ser Ala Arg Gln Arg Arg Glu Leu Arg Leu Phe Asp Ile Lys
80 85 90

```
Pro Glu Gln Gln Arg Tyr Ser Leu Phe Leu Pro Leu His Glu Leu
                 95                  100                 105
Trp Lys Gln Tyr Ile Arg Asp Leu Cys Ser Gly Leu Lys Pro Asp
                110                 115                 120
Thr Gln Pro Gln Met Ile Gln Ala Lys Leu Leu Lys Ala Asp Leu
                125                 130                 135
His Gly Ala Ile Ile Ser Val Thr Lys Ser Lys Cys Pro Ser Tyr
                140                 145                 150
Val Gly Ile Thr Gly Ile Leu Leu Gln Glu Thr Lys His Ile Phe
                155                 160                 165
Lys Ile Ile Thr Lys Glu Asp Arg Leu Lys Val Ile Pro Lys Leu
                170                 175                 180
Asn Cys Val Phe Thr Val Glu Thr Asp Gly Phe Ile Ser Tyr Ile
                185                 190                 195
Tyr Gly Ser Lys Phe Gln Leu Arg Ser Ser Glu Arg Ser Ala Lys
                200                 205                 210
Lys Phe Lys Ala Lys Gly Thr Ile Asp Leu
                215                 220


<210> 5
<211> 235
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1911910CD1

<400> 5
Met Gly Ser Thr Glu Ser Ser Glu Gly Arg Arg Val Ser Phe Gly
  1               5                  10                  15
Val Asp Glu Glu Glu Arg Val Arg Val Leu Gln Gly Val Arg Leu
                 20                  25                  30
Ser Glu Asn Val Val Asn Arg Met Lys Glu Pro Ser Ser Pro Pro
                 35                  40                  45
Pro Ala Pro Thr Ser Ser Thr Phe Gly Leu Gln Asp Gly Asn Leu
                 50                  55                  60
Arg Ala Pro His Lys Glu Ser Thr Leu Pro Arg Ser Gly Ser Ser
                 65                  70                  75
Gly Gly Gln Gln Pro Ser Gly Met Lys Glu Gly Val Lys Arg Tyr
                 80                  85                  90
Glu Gln Glu His Ala Ala Ile Gln Asp Lys Leu Phe Gln Val Ala
                 95                 100                 105
Lys Arg Glu Arg Glu Ala Ala Thr Lys His Ser Lys Ala Ser Leu
                110                 115                 120
Pro Thr Gly Glu Gly Ser Ile Ser His Glu Glu Gln Lys Ser Val
                125                 130                 135
Arg Leu Ala Arg Glu Leu Glu Ser Arg Glu Ala Glu Leu Arg Arg
                140                 145                 150
Arg Asp Thr Phe Tyr Lys Glu Gln Leu Glu Arg Ile Glu Arg Lys
                155                 160                 165
Asn Ala Glu Met Tyr Lys Leu Ser Ser Glu Gln Phe His Glu Ala
                170                 175                 180
Ala Ser Lys Met Glu Ser Thr Ile Lys Pro Arg Arg Val Glu Pro
                185                 190                 195
Val Cys Ser Gly Leu Gln Ala Gln Ile Leu His Cys Tyr Arg Asp
                200                 205                 210
Arg Pro His Glu Val Leu Leu Cys Ser Asp Leu Val Lys Ala Tyr
                215                 220                 225
Gln Arg Cys Val Ser Ala Ala His Lys Gly
                230                 235


<210> 6
<211> 487
<212> PRT
```

<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1928920CD1

<400> 6
```
Met Ala Ser Ser Ala Glu Gly Asp Glu Gly Thr Val Val Ala Leu
 1               5                  10                  15
Ala Gly Val Leu Gln Ser Gly Phe Gln Glu Leu Ser Leu Asn Lys
                20                  25                  30
Leu Ala Thr Ser Leu Gly Ala Ser Glu Gln Ala Leu Arg Leu Ile
                35                  40                  45
Ile Ser Ile Phe Leu Gly Tyr Pro Phe Ala Leu Phe Tyr Arg His
                50                  55                  60
Tyr Leu Phe Tyr Lys Glu Thr Tyr Leu Ile His Leu Phe His Thr
                65                  70                  75
Phe Thr Gly Leu Ser Ile Ala Tyr Phe Asn Phe Gly Asn Gln Leu
                80                  85                  90
Tyr His Ser Leu Leu Cys Ile Val Leu Gln Phe Leu Ile Leu Arg
                95                 100                 105
Leu Met Gly Arg Thr Ile Thr Ala Val Leu Thr Thr Phe Cys Phe
               110                 115                 120
Gln Met Ala Tyr Leu Leu Ala Gly Tyr Tyr Tyr Thr Ala Thr Gly
               125                 130                 135
Asn Tyr Asp Ile Lys Trp Thr Met Pro His Cys Val Leu Thr Leu
               140                 145                 150
Lys Leu Ile Gly Leu Ala Val Asp Tyr Phe Asp Gly Gly Lys Asp
               155                 160                 165
Gln Asn Ser Leu Ser Ser Glu Gln Gln Lys Tyr Ala Ile Arg Gly
               170                 175                 180
Val Pro Ser Leu Leu Glu Val Ala Gly Phe Ser Tyr Phe Tyr Gly
               185                 190                 195
Ala Phe Leu Val Gly Pro Gln Phe Ser Met Asn His Tyr Met Lys
               200                 205                 210
Leu Val Gln Gly Glu Leu Ile Asp Ile Pro Gly Lys Ile Pro Asn
               215                 220                 225
Ser Ile Ile Pro Ala Leu Lys Arg Leu Ser Leu Gly Leu Phe Tyr
               230                 235                 240
Leu Val Gly Tyr Thr Leu Leu Ser Pro His Ile Thr Glu Asp Tyr
               245                 250                 255
Leu Leu Thr Glu Asp Tyr Asp Asn His Pro Phe Trp Phe Arg Cys
               260                 265                 270
Met Tyr Met Leu Ile Trp Gly Lys Phe Val Leu Tyr Lys Tyr Val
               275                 280                 285
Thr Cys Trp Leu Val Thr Glu Gly Val Cys Ile Leu Thr Gly Leu
               290                 295                 300
Gly Phe Asn Gly Phe Glu Glu Lys Gly Lys Ala Lys Trp Asp Ala
               305                 310                 315
Cys Ala Asn Met Lys Val Trp Leu Phe Glu Thr Asn Pro Arg Phe
               320                 325                 330
Thr Gly Thr Ile Ala Ser Phe Asn Ile Asn Thr Asn Ala Trp Val
               335                 340                 345
Ala Arg Tyr Ile Phe Lys Arg Leu Lys Phe Leu Gly Asn Lys Glu
               350                 355                 360
Leu Ser Gln Gly Leu Ser Leu Leu Phe Leu Ala Leu Trp His Gly
               365                 370                 375
Leu His Ser Gly Tyr Leu Val Cys Phe Gln Met Glu Phe Leu Ile
               380                 385                 390
Val Ile Val Glu Arg Gln Ala Ala Arg Leu Ile Gln Glu Ser Pro
               395                 400                 405
Thr Leu Ser Lys Leu Ala Ala Ile Thr Val Leu Gln Pro Phe Tyr
               410                 415                 420
Tyr Leu Val Gln Gln Thr Ile His Trp Leu Phe Met Gly Tyr Ser
               425                 430                 435
Met Thr Ala Phe Cys Leu Phe Thr Trp Asp Lys Trp Leu Lys Val
               440                 445                 450
```

```
Tyr Lys Ser Ile Tyr Phe Leu Gly His Ile Phe Phe Leu Ser Leu
                455                 460                 465
Leu Phe Ile Leu Pro Tyr Ile His Lys Ala Met Val Pro Arg Lys
                470                 475                 480
Glu Lys Leu Lys Lys Met Glu
                485


<210> 7
<211> 212
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2170846CD1

<400> 7
Met Ala Ala Pro Pro Gln Leu Arg Ala Leu Leu Val Val Val Asn
 1               5                  10                  15
Ala Leu Leu Arg Lys Arg Arg Tyr His Ala Ala Leu Ala Val Leu
                20                  25                  30
Lys Gly Phe Arg Asn Gly Ala Val Tyr Gly Ala Lys Ile Arg Ala
                35                  40                  45
Pro His Ala Leu Val Met Thr Phe Leu Phe Arg Asn Gly Ser Leu
                50                  55                  60
Gln Glu Lys Leu Trp Ala Ile Leu Gln Ala Thr Tyr Ile His Ser
                65                  70                  75
Trp Asn Leu Ala Arg Phe Val Phe Thr Tyr Lys Gly Leu Arg Ala
                80                  85                  90
Leu Gln Ser Tyr Ile Gln Gly Lys Thr Tyr Pro Ala His Ala Phe
                95                 100                 105
Leu Ala Ala Phe Leu Gly Gly Ile Leu Val Phe Gly Glu Asn Asn
               110                 115                 120
Asn Ile Asn Ser Gln Ile Asn Met Tyr Leu Leu Ser Arg Val Leu
               125                 130                 135
Phe Ala Leu Ser Arg Leu Ala Val Glu Lys Gly Tyr Ile Pro Glu
               140                 145                 150
Pro Arg Trp Asp Pro Phe Pro Leu Leu Thr Ala Val Val Trp Gly
               155                 160                 165
Leu Val Leu Trp Leu Phe Glu Tyr His Arg Ser Thr Leu Gln Pro
               170                 175                 180
Ser Leu Gln Ser Ser Met Thr Tyr Leu Tyr Glu Asp Ser Asn Val
               185                 190                 195
Trp His Asp Ile Ser Asp Phe Leu Ile Tyr Asn Lys Ser Arg Pro
               200                 205                 210
Ser Asn


<210> 8
<211> 241
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2176361CD1

<400> 8
Met Ala Pro Val Arg Arg Ser Ala Lys Trp Arg Pro Gly Gly Ile
 1               5                  10                  15
Glu Ala Arg Gly Glu Gly Val Ser Thr Val Gly Tyr Arg Asn Lys
                20                  25                  30
Asn Val Arg Gln Lys Thr Trp Arg Pro Asn His Pro Gln Ala Phe
                35                  40                  45
Val Gly Ser Val Arg Glu Gly Gln Gly Phe Ala Phe Arg Arg Lys
                50                  55                  60
```

```
Leu Lys Ile Gln Gln Ser Tyr Lys Lys Leu Leu Arg Lys Glu Lys
                65                  70                  75
Lys Ala Gln Thr Ser Leu Glu Ser Gln Phe Thr Asp Arg Tyr Pro
                80                  85                  90
Asp Asn Leu Lys His Leu Tyr Leu Ala Glu Glu Glu Arg His Arg
                95                  100                 105
Lys Gln Ala Arg Lys Val Asp His Pro Leu Ser Glu Gln Val His
                110                 115                 120
Gln Pro Leu Leu Glu Glu Gln Cys Ser Ile Asp Glu Pro Leu Phe
                125                 130                 135
Glu Asp Gln Cys Ser Phe Asp Gln Pro Gln Pro Glu Glu Gln Cys
                140                 145                 150
Ile Lys Thr Val Asn Ser Phe Thr Ile Pro Lys Lys Asn Lys Lys
                155                 160                 165
Lys Thr Ser Asn Gln Lys Ala Gln Glu Glu Tyr Glu Gln Ile Gln
                170                 175                 180
Ala Lys Arg Ala Ala Lys Lys Gln Glu Phe Glu Arg Arg Lys Gln
                185                 190                 195
Glu Arg Glu Glu Ala Gln Arg Gln Tyr Lys Lys Lys Lys Met Glu
                200                 205                 210
Val Phe Lys Ile Leu Asn Lys Lys Thr Lys Lys Gly Gln Pro Asn
                215                 220                 225
Leu Asn Val Gln Met Glu Tyr Leu Leu Gln Lys Ile Gln Glu Lys
                230                 235                 240
Cys


<210> 9
<211> 375
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2212732CD1

<400> 9
Met Pro Gln Glu Leu Pro Gln Ser Pro Arg Thr Arg Gln Pro Glu
 1               5                  10                  15
Pro Asp Phe Tyr Cys Val Lys Trp Ile Pro Trp Lys Gly Glu Gln
                20                  25                  30
Thr Pro Ile Ile Thr Gln Ser Thr Asn Gly Pro Cys Pro Leu Leu
                35                  40                  45
Ala Ile Met Asn Ile Leu Phe Leu Gln Trp Lys Val Lys Leu Pro
                50                  55                  60
Pro Gln Lys Glu Val Ile Thr Ser Asp Glu Leu Met Ala His Leu
                65                  70                  75
Gly Asn Cys Leu Leu Ser Ile Lys Pro Gln Glu Lys Ser Glu Gly
                80                  85                  90
Leu Gln Leu Asn Phe Gln Gln Asn Val Asp Asp Ala Met Thr Val
                95                  100                 105
Leu Pro Lys Leu Ala Thr Gly Leu Asp Val Asn Val Arg Phe Thr
                110                 115                 120
Gly Val Ser Asp Phe Glu Tyr Thr Pro Glu Cys Ser Val Phe Asp
                125                 130                 135
Leu Leu Gly Ile Pro Leu Tyr His Gly Trp Leu Val Asp Pro Gln
                140                 145                 150
Gln Ser Pro Glu Ala Val Arg Ala Val Gly Lys Leu Ser Tyr Asn
                155                 160                 165
Gln Leu Val Glu Arg Ile Ile Thr Cys Lys His Ser Ser Asp Thr
                170                 175                 180
Asn Leu Val Thr Glu Gly Leu Ile Ala Glu Gln Phe Leu Glu Thr
                185                 190                 195
Thr Ala Ala Gln Leu Thr Tyr His Gly Leu Cys Glu Leu Thr Ala
                200                 205                 210
Ala Ala Lys Glu Gly Glu Leu Ser Val Phe Phe Arg Asn Asn His
                215                 220                 225
```

Phe Ser Thr Met Thr Lys His Lys Ser His Leu Tyr Leu Leu Val
230 235 240
Thr Asp Gln Gly Phe Leu Gln Glu Glu Gln Val Val Trp Glu Ser
245 250 255
Leu His Asn Val Asp Gly Asp Ser Cys Phe Cys Asp Ser Asp Phe
260 265 270
His Leu Ser His Ser Leu Gly Lys Gly Pro Gly Ala Glu Gly Gly
275 280 285
Ser Gly Ser Pro Glu Lys Gln Leu Gln Val Asp Gln Asp Tyr Leu
290 295 300
Ile Ala Leu Ser Leu Gln Gln Gln Gln Pro Arg Gly Pro Leu Gly
305 310 315
Leu Thr Asp Leu Glu Leu Ala Gln Gln Leu Gln Gln Glu Glu Tyr
320 325 330
Gln Gln Gln Gln Ala Ala Gln Pro Val Arg Met Arg Thr Arg Val
335 340 345
Leu Ser Leu Gln Gly Arg Gly Ala Thr Ser Gly Arg Pro Ala Gly
350 355 360
Glu Arg Arg Gln Arg Pro Lys His Glu Ser Asp Cys Ile Leu Leu
365 370 375

<210> 10
<211> 429
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2303457CD1

<400> 10
Met Ser Asn Arg Asn Asn Asn Lys Leu Pro Ser Asn Leu Pro Gln
1 5 10 15
Leu Gln Asn Leu Ile Lys Arg Asp Pro Pro Ala Tyr Ile Glu Glu
20 25 30
Phe Leu Gln Gln Tyr Asn His Tyr Lys Ser Asn Val Glu Ile Phe
35 40 45
Lys Leu Gln Pro Asn Lys Pro Ser Lys Glu Leu Ala Glu Leu Val
50 55 60
Met Phe Met Ala Gln Ile Ser His Cys Tyr Pro Glu Tyr Leu Ser
65 70 75
Asn Phe Pro Gln Glu Val Lys Asp Leu Leu Ser Cys Asn His Thr
80 85 90
Val Leu Asp Pro Asp Leu Arg Met Thr Phe Cys Lys Ala Leu Ile
95 100 105
Leu Leu Arg Asn Lys Asn Leu Ile Asn Pro Ser Ser Leu Leu Glu
110 115 120
Leu Phe Phe Glu Leu Phe Arg Cys His Asp Lys Leu Leu Arg Lys
125 130 135
Thr Leu Tyr Thr His Ile Val Thr Asp Ile Lys Asn Ile Asn Ala
140 145 150
Lys His Lys Asn Asn Lys Val Asn Val Val Leu Gln Asn Phe Met
155 160 165
Tyr Thr Met Leu Arg Asp Ser Asn Ala Thr Ala Ala Lys Met Ser
170 175 180
Leu Asp Val Met Ile Glu Leu Tyr Arg Arg Asn Ile Trp Asn Asp
185 190 195
Ala Lys Thr Val Asn Val Ile Thr Thr Ala Cys Phe Ser Lys Val
200 205 210
Thr Lys Ile Leu Val Ala Ala Leu Thr Phe Phe Leu Gly Lys Asp
215 220 225
Glu Asp Glu Lys Gln Asp Ser Asp Ser Glu Ser Glu Asp Asp Gly
230 235 240
Pro Thr Ala Arg Asp Leu Leu Val Gln Tyr Ala Thr Gly Lys Lys
245 250 255
Ser Ser Lys Asn Lys Lys Lys Leu Glu Lys Ala Met Lys Val Leu

```
                    260                     265                 270
Lys Lys Gln Lys Lys Lys Lys Lys Pro Glu Val Phe Asn Phe Ser
                    275                     280                 285
Ala Ile His Leu Ile His Asp Pro Gln Asp Phe Ala Glu Lys Leu
                    290                     295                 300
Leu Lys Gln Leu Glu Cys Cys Lys Glu Arg Phe Glu Val Lys Met
                    305                     310                 315
Met Leu Met Asn Leu Ile Ser Arg Leu Val Gly Ile His Glu Leu
                    320                     325                 330
Phe Leu Phe Asn Phe Tyr Pro Phe Leu Lys Arg Phe Leu Lys Pro
                    335                     340                 345
His Gln Arg Glu Val Thr Lys Ile Leu Leu Phe Val Glu Lys Asp
                    350                     355                 360
Ser His His Leu Val Pro Gln Gly Phe Phe Asn Ser Trp Leu Met
                    365                     370                 375
Leu Gly Glu Lys Ile Phe Phe Asn Gly Lys Lys Ser Gly Lys Met
                    380                     385                 390
Leu Met Thr Val Gly Asn Leu Met Val Lys Arg Gly Val Tyr Lys
                    395                     400                 405
Arg Ser Lys Val Phe Leu Gly Gly Asn Ser Val Gly Arg Asn Phe
                    410                     415                 420
Phe Gln Lys Asn Pro Gly Gly Ser Ser
                    425
```

<210> 11
<211> 329
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2317552CD1

<400> 11

```
Met Glu Val Ala Glu Pro Ser Ser Pro Thr Glu Glu Glu Glu Glu
  1                 5                    10                  15
Glu Glu Glu His Ser Ala Glu Pro Arg Pro Arg Thr Arg Ser Asn
                    20                   25                  30
Pro Glu Gly Ala Glu Asp Arg Ala Val Gly Ala Gln Ala Ser Val
                    35                   40                  45
Gly Ser Arg Ser Glu Gly Glu Gly Glu Ala Ala Ser Ala Asp Asp
                    50                   55                  60
Gly Ser Leu Asn Thr Ser Gly Ala Gly Pro Lys Ser Trp Gln Val
                    65                   70                  75
Pro Pro Pro Ala Pro Glu Val Gln Ile Arg Thr Pro Arg Val Asn
                    80                   85                  90
Cys Pro Glu Lys Val Ile Ile Cys Leu Asp Leu Ser Glu Glu Met
                    95                   100                 105
Ser Leu Pro Lys Leu Glu Ser Phe Asn Gly Ser Lys Thr Asn Ala
                    110                  115                 120
Leu Asn Val Ser Gln Lys Met Ile Glu Met Phe Val Arg Thr Lys
                    125                  130                 135
His Lys Ile Asp Lys Ser His Glu Phe Ala Leu Val Val Val Asn
                    140                  145                 150
Asp Asp Thr Ala Trp Leu Ser Gly Leu Thr Ser Asp Pro Arg Glu
                    155                  160                 165
Leu Cys Ser Cys Leu Tyr Asp Leu Glu Thr Ala Ser Cys Ser Thr
                    170                  175                 180
Phe Asn Leu Glu Gly Leu Phe Ser Leu Ile Gln Gln Lys Thr Glu
                    185                  190                 195
Leu Pro Val Thr Glu Asn Val Gln Thr Ile Pro Pro Pro Tyr Val
                    200                  205                 210
Val Arg Thr Ile Leu Val Tyr Ser Arg Pro Pro Cys Gln Pro Gln
                    215                  220                 225
Phe Ser Leu Thr Glu Pro Met Lys Lys Met Phe Gln Cys Pro Tyr
                    230                  235                 240
```

```
Phe Phe Phe Asp Val Val Tyr Ile His Asn Gly Thr Glu Glu Lys
                245                 250                 255
Glu Glu Glu Met Ser Trp Lys Asp Met Phe Ala Phe Met Gly Ser
                260                 265                 270
Leu Asp Thr Lys Gly Thr Ser Tyr Lys Tyr Glu Val Ala Leu Ala
                275                 280                 285
Gly Pro Ala Leu Glu Leu His Asn Cys Met Ala Lys Leu Leu Ala
                290                 295                 300
His Pro Leu Gln Arg Pro Cys Gln Ser His Ala Ser Tyr Ser Leu
                305                 310                 315
Leu Glu Glu Glu Asp Glu Ala Ile Glu Val Glu Ala Thr Val
                320                 325


<210> 12
<211> 476
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2416366CD1

<400> 12
Met Gln Asn Asp Ser Phe His Ser Asp Ser His Met Asp Arg Lys
  1               5                  10                  15
Lys Phe His Ser Ser Asp Ser Glu Glu Glu Glu His Lys Lys Gln
                 20                  25                  30
Lys Met Asp Ser Asp Glu Asp Glu Lys Glu Gly Glu Glu Glu Lys
                 35                  40                  45
Val Ala Lys Arg Lys Ala Ala Val Leu Ser Asp Ser Glu Asp Glu
                 50                  55                  60
Glu Lys Ala Ser Ala Lys Lys Ser Arg Val Val Ser Asp Ala Asp
                 65                  70                  75
Asp Ser Asp Ser Asp Ala Val Ser Asp Lys Ser Gly Lys Arg Glu
                 80                  85                  90
Lys Thr Ile Ala Ser Asp Ser Glu Glu Glu Ala Gly Lys Glu Leu
                 95                 100                 105
Ser Asp Lys Lys Asn Glu Glu Lys Asp Leu Phe Gly Ser Asp Ser
                110                 115                 120
Glu Ser Gly Asn Glu Glu Glu Asn Leu Ile Ala Asp Ile Phe Gly
                125                 130                 135
Glu Ser Gly Asp Glu Glu Glu Glu Glu Phe Thr Gly Phe Asn Gln
                140                 145                 150
Glu Asp Leu Glu Glu Glu Lys Gly Glu Thr Gln Val Lys Glu Ala
                155                 160                 165
Glu Asp Ser Asp Ser Asp Asp Asn Ile Lys Arg Gly Lys His Met
                170                 175                 180
Asp Phe Leu Ser Asp Phe Glu Met Met Leu Gln Arg Lys Lys Ser
                185                 190                 195
Met Ser Gly Lys Arg Arg Arg Asn Arg Asp Gly Gly Thr Phe Ile
                200                 205                 210
Ser Asp Ala Asp Asp Val Val Ser Ala Met Ile Val Lys Met Asn
                215                 220                 225
Glu Ala Ala Glu Glu Asp Arg Gln Leu Asn Asn Gln Lys Lys Pro
                230                 235                 240
Ala Leu Lys Lys Leu Thr Leu Leu Pro Ala Val Val Met His Leu
                245                 250                 255
Lys Lys Gln Asp Leu Lys Glu Thr Phe Ile Asp Ser Gly Val Met
                260                 265                 270
Ser Ala Ile Lys Glu Trp Leu Ser Pro Leu Pro Asp Arg Ser Leu
                275                 280                 285
Pro Ala Leu Lys Ile Arg Glu Glu Leu Leu Lys Ile Leu Gln Glu
                290                 295                 300
Leu Pro Ser Val Ser Gln Glu Thr Leu Lys His Ser Gly Ile Gly
                305                 310                 315
Arg Ala Val Met Tyr Leu Tyr Lys His Pro Lys Glu Ser Arg Ser
```


WEST

```
                320                             325                             330
Asn Lys Asp Met Ala Gly Lys Leu Ile Asn Glu Trp Ser Arg Pro
                335                             340                             345
Ile Phe Gly Leu Thr Ser Asn Tyr Lys Gly Met Thr Arg Glu Glu
                350                             355                             360
Arg Glu Gln Arg Asp Leu Glu Gln Met Pro Gln Arg Arg Arg Met
                365                             370                             375
Asn Ser Thr Gly Gly Gln Thr Pro Arg Arg Asp Leu Glu Lys Val
                380                             385                             390
Leu Thr Gly Glu Glu Lys Ala Leu Arg Pro Gly Asp Pro Gly Phe
                395                             400                             405
Cys Ala Arg Ala Arg Val Pro Met Pro Ser Asn Lys Asp Tyr Val
                410                             415                             420
Val Arg Pro Lys Trp Asn Val Glu Met Glu Ser Ser Arg Phe Gln
                425                             430                             435
Ala Thr Ser Lys Lys Gly Ile Ser Arg Leu Asp Lys Gln Met Arg
                440                             445                             450
Lys Phe Thr Asp Ile Arg Lys Lys Ser Arg Ser Ala His Ala Val
                455                             460                             465
Lys Ile Ser Ile Glu Gly Asn Lys Met Pro Leu
                470                             475
```

<210> 13
<211> 366
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2472980CD1

<400> 13

```
Met Ala Ala Ala Tyr Phe Pro Asp Cys Ile Val Arg Pro Phe Gly
 1               5                              10                              15
Ser Ser Val Asn Thr Phe Gly Lys Leu Gly Cys Asp Leu Asp Met
                20                              25                              30
Phe Leu Asp Leu Asp Glu Thr Arg Asn Leu Ser Ala His Lys Ile
                35                              40                              45
Ser Gly Asn Phe Leu Met Glu Phe Gln Val Lys Asn Val Pro Ser
                50                              55                              60
Glu Arg Ile Ala Thr Gln Lys Ile Leu Ser Val Leu Gly Glu Cys
                65                              70                              75
Leu Asp His Phe Gly Pro Gly Cys Val Gly Val Gln Lys Ile Leu
                80                              85                              90
Asn Ala Arg Cys Pro Leu Val Arg Phe Ser His Gln Ala Ser Gly
                95                              100                             105
Phe Gln Cys Asp Leu Thr Thr Asn Asn Arg Ile Ala Leu Thr Ser
                110                             115                             120
Ser Glu Leu Leu Tyr Ile Tyr Gly Ala Leu Asp Ser Arg Val Arg
                125                             130                             135
Ala Leu Val Phe Ser Val Arg Cys Trp Ala Arg Ala His Ser Leu
                140                             145                             150
Thr Ser Ser Ile Pro Gly Ala Trp Ile Thr Asn Phe Ser Leu Thr
                155                             160                             165
Met Met Val Ile Phe Phe Leu Gln Arg Arg Ser Pro Pro Ile Leu
                170                             175                             180
Pro Thr Leu Asp Ser Leu Lys Thr Leu Ala Asp Ala Glu Asp Lys
                185                             190                             195
Cys Val Ile Glu Gly Asn Asn Cys Thr Phe Val Arg Asp Leu Ser
                200                             205                             210
Arg Ile Lys Pro Ser Gln Asn Thr Glu Thr Leu Glu Leu Leu Leu
                215                             220                             225
Lys Glu Phe Phe Glu Tyr Phe Gly Asn Phe Ala Phe Asp Lys Asn
                230                             235                             240
Ser Ile Asn Ile Arg Gln Gly Arg Glu Gln Asn Lys Pro Asp Ser
                245                             250                             255
```


WEST

```
Ser Pro Leu Tyr Ile Gln Asn Pro Phe Glu Thr Ser Leu Asn Ile
                260                 265                 270
Ser Lys Asn Val Ser Gln Ser Gln Leu Gln Lys Phe Val Asp Leu
                275                 280                 285
Ala Arg Glu Ser Ala Trp Ile Leu Gln Gln Glu Asp Thr Asp Arg
                290                 295                 300
Pro Ser Ile Ser Ser Asn Arg Pro Trp Gly Leu Val Ser Leu Leu
                305                 310                 315
Leu Pro Ser Ala Pro Asn Arg Lys Ser Phe Thr Lys Lys Lys Ser
                320                 325                 330
Asn Lys Phe Ala Ile Glu Thr Val Lys Asn Leu Leu Glu Ser Leu
                335                 340                 345
Lys Gly Asn Arg Thr Glu Asn Phe Thr Lys Thr Ser Gly Lys Arg
                350                 355                 360
Thr Ile Ser Thr Gln Thr
                365


<210> 14
<211> 152
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2541640CD1

<400> 14
Met Gly Gly Val Gly Val Ala Glu Ala Ala Arg Pro Leu Leu Ser
  1             5                   10                  15
Trp Pro Thr Ile Ser Leu Thr Ile Phe Thr Ala Val Asn Ser Ser
                20                  25                  30
Gln Gly Gly Gly Leu Val Gln Arg Gln Leu Arg Phe His Asn Ser
                35                  40                  45
His Arg Val Leu Cys Arg Arg Cys Pro Cys Pro Pro Thr Pro Ala
                50                  55                  60
Trp Trp Glu Cys Asp Ala Arg Leu Leu Pro Pro Pro Trp Pro Pro
                65                  70                  75
Val Pro Pro Ala Ser Thr Ser Pro Glu Ile Leu Pro Thr Pro His
                80                  85                  90
Leu His Arg Ser Pro His Ala Pro Gly Ala Pro Lys Pro Pro Pro
                95                  100                 105
Asn Pro Thr His Pro Gly Ala Gly Thr Gly Val Ser Glu Leu Ser
                110                 115                 120
Gln Gly Pro Trp Glu Val Ala Gly Thr Gly Ala Ser Cys Ser Leu
                125                 130                 135
Phe His Phe Pro Phe Arg Ile Trp Pro Gly Trp Arg Thr Gly Gln
                140                 145                 150
Asp Gly


<210> 15
<211> 233
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2695204CD1

<400> 15
Met Gly Arg Arg Leu Lys Gly Ala Arg Arg Leu Lys Leu Ser Pro
  1             5                   10                  15
Leu Arg Ser Leu Arg Lys Gly Pro Gly Leu Leu Ser Pro Pro Ser
                20                  25                  30
Ala Ser Pro Val Pro Thr Pro Ala Val Ser Arg Thr Leu Leu Gly
                35                  40                  45
```

```
Asn Phe Glu Glu Ser Leu Leu Arg Gly Arg Phe Ala Pro Ser Gly
                50                  55                  60
His Ile Glu Gly Phe Thr Ala Glu Ile Gly Ala Ser Gly Ser Tyr
                65                  70                  75
Cys Pro Gln His Val Thr Leu Pro Val Thr Val Thr Phe Phe Asp
                80                  85                  90
Val Ser Glu Gln Asn Ala Pro Ala Pro Phe Leu Gly Ile Val Asp
                95                 100                 105
Leu Asn Pro Leu Gly Arg Lys Gly Tyr Ser Val Pro Lys Val Gly
               110                 115                 120
Thr Val Gln Val Thr Leu Phe Asn Pro Asn Gln Thr Val Val Lys
               125                 130                 135
Met Phe Leu Val Thr Phe Asp Phe Ser Asp Met Pro Ala Ala His
               140                 145                 150
Met Thr Phe Leu Arg His Arg Leu Phe Leu Val Pro Val Gly Glu
               155                 160                 165
Glu Gly Asn Ala Asn Pro Thr His Arg Leu Leu Cys Tyr Leu Leu
               170                 175                 180
His Leu Arg Phe Arg Ser Ser Arg Ser Gly Arg Leu Ser Leu His
               185                 190                 195
Gly Asp Ile Arg Leu Leu Phe Ser Arg Arg Ser Leu Glu Leu Asp
               200                 205                 210
Thr Gly Leu Pro Tyr Glu Leu Gln Ala Val Thr Glu Ala Pro His
               215                 220                 225
Asn Pro Arg Tyr Ser Pro Leu Pro
               230
```

<210> 16
<211> 357
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2805526CD1

<400> 16

```
Met Glu Val Leu Arg Pro Gln Leu Ile Arg Ile Asp Gly Arg Asn
  1               5                  10                  15
Tyr Arg Lys Asn Pro Val Gln Glu Gln Thr Tyr Gln His Glu Glu
                 20                  25                  30
Asp Glu Glu Asp Phe Tyr Gln Gly Ser Met Glu Cys Ala Asp Glu
                 35                  40                  45
Pro Cys Asp Ala Tyr Glu Val Glu Gln Thr Pro Gln Gly Phe Arg
                 50                  55                  60
Ser Thr Leu Arg Ala Pro Ser Leu Leu Tyr Lys His Ile Val Gly
                 65                  70                  75
Lys Arg Gly Asp Thr Arg Lys Lys Ile Glu Met Glu Thr Lys Thr
                 80                  85                  90
Ser Ile Ser Ile Pro Lys Pro Gly Gln Asp Gly Glu Ile Val Ile
                 95                 100                 105
Thr Gly Gln His Arg Asn Gly Val Ile Ser Ala Arg Thr Arg Ile
                110                 115                 120
Asp Val Leu Leu Asp Thr Phe Arg Arg Lys Gln Pro Phe Thr His
                125                 130                 135
Phe Leu Ala Phe Phe Leu Asn Glu Val Glu Val Gln Glu Gly Phe
                140                 145                 150
Leu Arg Phe Gln Glu Glu Val Leu Ala Lys Cys Ser Met Asp His
                155                 160                 165
Gly Val Asp Ser Ser Ile Phe Gln Asn Pro Lys Lys Leu His Leu
                170                 175                 180
Thr Ile Gly Met Leu Val Leu Leu Ser Glu Glu Glu Ile Gln Gln
                185                 190                 195
Thr Cys Glu Met Leu Gln Gln Cys Lys Glu Glu Phe Ile Asn Asp
                200                 205                 210
Ile Ser Gly Gly Lys Pro Leu Glu Val Glu Met Ala Gly Ile Glu
```

```
                215                     220                     225
Tyr Met Asn Asp Asp Pro Gly Met Val Asp Val Leu Tyr Ala Lys
                230                     235                     240
Val His Met Lys Asp Gly Ser Asn Arg Leu Gln Glu Leu Val Asp
                245                     250                     255
Arg Val Leu Glu Arg Phe Gln Ala Ser Gly Leu Ile Val Lys Glu
                260                     265                     270
Trp Asn Ser Val Lys Leu His Ala Thr Val Met Asn Thr Leu Phe
                275                     280                     285
Arg Lys Asp Pro Asn Ala Glu Gly Arg Tyr Asn Leu Tyr Thr Ala
                290                     295                     300
Glu Gly Lys Tyr Ile Phe Lys Glu Arg Glu Ser Phe Asp Gly Arg
                305                     310                     315
Asn Ile Leu Lys Leu Phe Glu Asn Phe Tyr Phe Gly Ser Leu Lys
                320                     325                     330
Leu Asn Ser Ile His Ile Ser Gln Arg Phe Thr Val Asp Ser Phe
                335                     340                     345
Gly Asn Tyr Ala Ser Cys Gly Gln Ile Asp Phe Ser
                350                     355
```

<210> 17
<211> 251
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2850382CD1

<400> 17

```
Met Glu Pro Gly Glu Glu Leu Glu Glu Glu Gly Ser Pro Gly Gly
 1               5                  10                  15
Arg Glu Asp Gly Phe Thr Ala Glu His Leu Ala Ala Glu Ala Met
                20                  25                  30
Ala Ala Asp Met Asp Pro Trp Leu Val Phe Asp Ala Arg Thr Thr
                35                  40                  45
Pro Ala Thr Glu Leu Asp Ala Trp Leu Ala Lys Tyr Pro Pro Ser
                50                  55                  60
Gln Val Thr Arg Tyr Gly Asp Pro Gly Ser Pro Asn Ser Glu Pro
                65                  70                  75
Val Gly Trp Ile Ala Val Tyr Gly Gln Gly Tyr Ser Pro Asn Ser
                80                  85                  90
Gly Asp Val Gln Gly Leu Gln Ala Ala Trp Glu Ala Leu Gln Thr
                95                 100                 105
Ser Gly Arg Pro Ile Thr Pro Gly Thr Leu Arg Gln Leu Ala Ile
               110                 115                 120
Thr His His Val Leu Ser Gly Lys Trp Leu Met His Leu Ala Pro
               125                 130                 135
Gly Phe Lys Leu Asp His Ala Trp Ala Gly Ile Ala Arg Ala Val
               140                 145                 150
Val Glu Gly Arg Leu Gln Val Ala Lys Val Ser Pro Arg Ala Lys
               155                 160                 165
Glu Gly Gly Arg Gln Val Ile Cys Val Tyr Thr Asp Asp Phe Thr
               170                 175                 180
Asp Arg Leu Gly Val Leu Glu Ala Asp Ser Ala Ile Arg Ala Ala
               185                 190                 195
Gly Ile Lys Cys Leu Leu Thr Tyr Lys Pro Asp Val Tyr Thr Tyr
               200                 205                 210
Leu Gly Ile Tyr Arg Ala Asn Arg Trp His Leu Cys Pro Thr Leu
               215                 220                 225
Tyr Glu Ser Arg Phe Gln Leu Gly Gly Ser Ala Arg Gly Ser Arg
               230                 235                 240
Val Leu Asp Arg Ala Asn Asn Val Glu Leu Thr
               245                 250
```

<210> 18
<211> 105
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2929276CD1

<400> 18
Met Ser Ile Tyr Phe Pro Ile His Cys Pro Asp Tyr Leu Arg Ser
1 5 10 15
Ala Lys Met Thr Glu Val Met Met Asn Thr Gln Pro Met Glu Glu
20 25 30
Ile Gly Leu Ser Pro Arg Lys Asp Gly Leu Ser Tyr Gln Ile Phe
35 40 45
Pro Asp Pro Ser Asp Phe Asp Arg Cys Cys Lys Leu Lys Asp Arg
50 55 60
Leu Pro Ser Ile Val Val Glu Pro Thr Glu Gly Glu Val Glu Ser
65 70 75
Gly Glu Leu Arg Trp Pro Pro Glu Glu Phe Leu Val Gln Glu Asp
80 85 90
Glu Gln Asp Asn Cys Glu Glu Thr Ala Lys Glu Asn Lys Glu Gln
95 100 105

<210> 19
<211> 876
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3033039CD1

<400> 19
Met Thr Met Asp Ala Leu Leu Ala Arg Leu Lys Leu Leu Asn Pro
1 5 10 15
Asp Asp Leu Arg Glu Glu Ile Val Lys Ala Gly Leu Lys Cys Gly
20 25 30
Pro Ile Thr Ser Thr Thr Arg Phe Ile Phe Glu Lys Lys Leu Ala
35 40 45
Gln Ala Leu Leu Glu Gln Gly Gly Arg Leu Ser Ser Phe Tyr His
50 55 60
His Glu Ala Gly Val Thr Ala Leu Ser Gln Asp Pro Gln Arg Ile
65 70 75
Leu Lys Pro Ala Glu Gly Asn Pro Thr Asp Gln Ala Gly Phe Ser
80 85 90
Glu Asp Arg Asp Phe Gly Tyr Ser Val Gly Leu Asn Pro Pro Glu
95 100 105
Glu Glu Ala Val Thr Ser Lys Thr Cys Ser Val Pro Pro Ser Asp
110 115 120
Thr Asp Thr Tyr Arg Ala Gly Ala Thr Ala Ser Lys Glu Pro Pro
125 130 135
Leu Tyr Tyr Gly Val Cys Pro Val Tyr Glu Asp Val Pro Ala Arg
140 145 150
Asn Glu Arg Ile Tyr Val Tyr Glu Asn Lys Lys Glu Ala Leu Gln
155 160 165
Ala Val Lys Met Ile Lys Gly Ser Arg Phe Lys Ala Phe Ser Thr
170 175 180
Arg Glu Asp Ala Glu Lys Phe Ala Arg Gly Ile Cys Asp Tyr Phe
185 190 195
Pro Ser Pro Ser Lys Thr Ser Leu Pro Leu Ser Pro Val Lys Thr
200 205 210
Ala Pro Leu Phe Ser Asn Asp Arg Leu Lys Asp Gly Leu Cys Leu
215 220 225
Ser Glu Ser Glu Thr Val Asn Lys Glu Arg Ala Asn Ser Tyr Lys

```
                230                     235                     240
Asn Pro Arg Thr Gln Asp Leu Thr Ala Lys Leu Arg Lys Ala Val
                245                     250                     255
Glu Lys Gly Glu Glu Asp Thr Phe Ser Asp Leu Ile Trp Ser Asn
                260                     265                     270
Pro Arg Tyr Leu Ile Gly Ser Gly Asp Asn Pro Thr Ile Val Gln
                275                     280                     285
Glu Gly Cys Arg Tyr Asn Val Met His Val Ala Ala Lys Glu Asn
                290                     295                     300
Gln Ala Ser Ile Cys Gln Leu Thr Leu Asp Val Leu Glu Asn Pro
                305                     310                     315
Asp Phe Met Arg Leu Met Tyr Pro Asp Asp Asp Glu Ala Met Leu
                320                     325                     330
Gln Lys Arg Ile Arg Tyr Val Val Asp Leu Tyr Leu Asn Thr Pro
                335                     340                     345
Asp Lys Met Gly Tyr Asp Thr Pro Leu His Phe Ala Cys Lys Phe
                350                     355                     360
Gly Asn Ala Asp Val Val Asn Val Leu Ser Ser His His Leu Ile
                365                     370                     375
Val Lys Asn Ser Arg Asn Lys Tyr Asp Lys Thr Pro Glu Asp Val
                380                     385                     390
Ile Cys Glu Arg Ser Lys Asn Lys Ser Val Glu Leu Lys Glu Arg
                395                     400                     405
Ile Arg Glu Tyr Leu Lys Gly His Tyr Tyr Val Pro Leu Leu Arg
                410                     415                     420
Ala Glu Glu Thr Ser Ser Pro Val Ile Gly Glu Leu Trp Ser Pro
                425                     430                     435
Asp Gln Thr Ala Glu Ala Ser His Val Ser Arg Tyr Gly Gly Ser
                440                     445                     450
Pro Arg Asp Pro Val Leu Thr Leu Arg Ala Phe Ala Gly Pro Leu
                455                     460                     465
Ser Pro Ala Lys Ala Glu Asp Phe Arg Lys Leu Trp Lys Thr Pro
                470                     475                     480
Pro Arg Glu Lys Ala Gly Phe Leu His His Val Lys Lys Ser Asp
                485                     490                     495
Pro Glu Arg Gly Phe Glu Arg Val Gly Arg Glu Leu Ala His Glu
                500                     505                     510
Leu Gly Tyr Pro Trp Val Glu Tyr Trp Glu Phe Leu Gly Cys Phe
                515                     520                     525
Val Asp Leu Ser Ser Gln Glu Gly Leu Gln Arg Leu Glu Glu Tyr
                530                     535                     540
Leu Thr Gln Gln Glu Ile Gly Lys Lys Ala Gln Gln Glu Thr Gly
                545                     550                     555
Glu Arg Glu Ala Ser Cys Arg Asp Lys Ala Thr Thr Ser Gly Ser
                560                     565                     570
Asn Ser Ile Ser Val Arg Ala Phe Leu Asp Glu Asp Asp Met Ser
                575                     580                     585
Leu Glu Glu Ile Lys Asn Arg Gln Asn Ala Ala Arg Asn Asn Ser
                590                     595                     600
Pro Pro Thr Val Gly Ala Phe Gly His Thr Arg Cys Ser Ala Phe
                605                     610                     615
Pro Leu Glu Gln Glu Ala Asp Leu Ile Glu Ala Ala Glu Pro Gly
                620                     625                     630
Gly Pro His Ser Ser Arg Asn Gly Leu Cys His Pro Leu Asn His
                635                     640                     645
Ser Arg Thr Leu Ala Gly Lys Arg Pro Lys Ala Pro Arg Gly Glu
                650                     655                     660
Glu Ala His Leu Pro Pro Val Ser Asp Leu Thr Val Glu Phe Asp
                665                     670                     675
Lys Leu Asn Leu Gln Asn Ile Gly Arg Ser Val Ser Lys Thr Pro
                680                     685                     690
Asp Glu Ser Thr Lys Thr Lys Asp Gln Ile Leu Thr Ser Arg Ile
                695                     700                     705
Asn Ala Val Glu Arg Asp Leu Leu Glu Pro Ser Pro Ala Asp Gln
                710                     715                     720
Leu Gly Asn Gly His Arg Arg Thr Glu Ser Glu Met Ser Ala Arg
                725                     730                     735
```

```
Ile Ala Lys Met Ser Leu Ser Pro Ser Ser Pro Arg His Glu Asp
                740                 745                 750
Gln Leu Glu Val Thr Arg Glu Pro Ala Arg Arg Leu Phe Leu Phe
                755                 760                 765
Gly Glu Glu Pro Ser Lys Leu Asp Gln Asp Val Leu Ala Ala Leu
                770                 775                 780
Glu Cys Ala Asp Val Asp Pro His Gln Phe Pro Ala Val His Arg
                785                 790                 795
Trp Lys Ser Ala Val Leu Cys Tyr Ser Pro Ser Asp Arg Gln Ser
                800                 805                 810
Trp Pro Ser Pro Ala Val Lys Gly Arg Phe Lys Ser Gln Leu Pro
                815                 820                 825
Asp Leu Ser Gly Pro His Ser Tyr Ser Pro Gly Arg Asn Ser Val
                830                 835                 840
Ala Gly Ser Asn Pro Ala Lys Pro Gly Leu Gly Ser Pro Gly Arg
                845                 850                 855
Tyr Ser Pro Val His Gly Ser Gln Leu Arg Arg Met Ala Arg Leu
                860                 865                 870
Ala Glu Leu Ala Ala Leu
                875
```

<210> 20
<211> 254
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3213216CD1

<400> 20

```
Met Leu Glu His Lys Ile Phe Lys Arg Phe Trp Leu Val Ile Asn
 1               5                  10                  15
Gln Glu Gly Asn Met Val Thr Ala Arg Gln Glu Pro Arg Leu Val
                20                  25                  30
Leu Ile Ser Leu Thr Cys Asp Gly Asp Thr Leu Thr Leu Ser Ala
                35                  40                  45
Ala Tyr Thr Lys Asp Leu Leu Leu Pro Ile Lys Thr Pro Thr Thr
                50                  55                  60
Asn Ala Val His Lys Cys Arg Val His Gly Leu Glu Ile Glu Gly
                65                  70                  75
Arg Asp Cys Gly Glu Ala Ala Ala Gln Trp Ile Thr Ser Phe Leu
                80                  85                  90
Lys Ser Gln Pro Tyr Arg Leu Val His Phe Glu Pro His Met Arg
                95                  100                 105
Pro Arg Arg Pro His Gln Ile Ala Asp Leu Phe Arg Pro Lys Asp
                110                 115                 120
Gln Ile Ala Tyr Ser Asp Thr Ser Pro Phe Leu Ile Leu Ser Glu
                125                 130                 135
Ala Ser Leu Ala Asp Leu Asn Ser Arg Leu Glu Lys Lys Val Lys
                140                 145                 150
Ala Thr Asn Phe Arg Pro Asn Ile Val Ile Ser Gly Cys Asp Val
                155                 160                 165
Tyr Ala Glu Asp Ser Trp Asp Glu Leu Leu Ile Gly Asp Val Glu
                170                 175                 180
Leu Lys Arg Val Met Ala Cys Ser Arg Cys Ile Leu Thr Thr Val
                185                 190                 195
Asp Pro Asp Thr Gly Val Met Ser Arg Lys Glu Pro Leu Glu Thr
                200                 205                 210
Leu Lys Ser Tyr Arg Gln Cys Asp Pro Ser Glu Arg Lys Leu Tyr
                215                 220                 225
Gly Lys Ser Pro Leu Phe Gly Gln Tyr Phe Val Leu Glu Asn Pro
                230                 235                 240
Gly Thr Ile Lys Val Gly Asp Pro Val Tyr Leu Leu Gly Gln
                245                 250
```


WEST

<210> 21
<211> 259
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3220944CD1

<400> 21
Met Lys Met Val Thr Gly Ala Val Ala Ser Val Leu Glu Asp Glu
1 5 10 15
Ala Thr Asp Thr Ser Asp Ser Glu Gly Ser Cys Gly Ser Glu Lys
20 25 30
Asp His Phe Tyr Ser Asp Asp Asp Ala Ile Glu Ala Asp Ser Glu
35 40 45
Gly Asp Ala Glu Pro Cys Asp Lys Glu Asn Glu Asn Asp Gly Glu
50 55 60
Ser Ser Val Gly Thr Asn Met Gly Trp Ala Asp Ala Met Ala Lys
65 70 75
Val Leu Asn Lys Lys Thr Pro Glu Ser Lys Pro Thr Ile Leu Val
80 85 90
Lys Asn Lys Lys Leu Glu Lys Glu Lys Glu Lys Leu Lys Gln Glu
95 100 105
Arg Leu Glu Lys Ile Lys Gln Arg Asp Lys Arg Leu Glu Trp Glu
110 115 120
Met Met Cys Arg Val Lys Pro Asp Val Val Gln Asp Lys Glu Thr
125 130 135
Glu Arg Asn Leu Gln Arg Ile Ala Thr Arg Gly Val Val Gln Leu
140 145 150
Phe Asn Ala Val Gln Lys His Gln Lys Asn Val Asp Glu Lys Val
155 160 165
Lys Glu Ala Gly Ser Ser Met Arg Lys Arg Ala Lys Leu Ile Ser
170 175 180
Thr Val Ser Lys Lys Asp Phe Ile Ser Val Leu Arg Gly Met Asp
185 190 195
Gly Ser Thr Asn Glu Thr Ala Ser Ser Arg Lys Lys Pro Lys Ala
200 205 210
Lys Gln Thr Glu Val Lys Ser Glu Glu Gly Pro Gly Trp Thr Ile
215 220 225
Leu Arg Asp Asp Phe Met Met Gly Ala Ser Met Lys Asp Trp Asp
230 235 240
Lys Glu Ser Asp Gly Pro Asp Asp Ser Arg Pro Glu Ser Ala Ser
245 250 255
Asp Ser Asp Thr

<210> 22
<211> 65
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3224631CD1

<400> 22
Met Asp Tyr Gly Val Gln Phe Thr Ala Ser Arg Lys Phe Leu Thr
1 5 10 15
Ile Thr Pro Ile Val Leu Tyr Phe Leu Thr Ser Phe Tyr Thr Lys
20 25 30
Tyr Asp Gln Ile His Phe Val Leu Asn Thr Val Ser Leu Met Ser
35 40 45
Val Leu Ile Pro Lys Leu Leu Gln Leu His Gly Val Arg Ile Phe
50 55 60
Gly Ile Asn Lys Tyr
65

<210> 23
<211> 163
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3242839CD1

<400> 23
```
Met Val Gly Gly Gly Gly Val Gly Gly Gly Leu Leu Glu Asn Ala
 1               5                  10                  15
Asn Pro Leu Ile Tyr Gln Arg Ser Gly Glu Arg Pro Val Thr Ala
                20                  25                  30
Gly Glu Glu Asp Glu Gln Val Pro Asp Ser Ile Asp Ala Arg Glu
                35                  40                  45
Ile Phe Asp Leu Ile Arg Ser Ile Asn Asp Pro Glu His Pro Leu
                50                  55                  60
Thr Leu Glu Glu Leu Asn Val Val Glu Gln Val Arg Val Gln Val
                65                  70                  75
Ser Asp Pro Glu Ser Thr Val Ala Val Ala Phe Thr Pro Thr Ile
                80                  85                  90
Pro His Cys Ser Met Ala Thr Leu Ile Gly Leu Ser Ile Lys Val
                95                 100                 105
Lys Leu Leu Arg Ser Leu Pro Gln Arg Phe Lys Met Asp Val His
               110                 115                 120
Ile Thr Pro Gly Thr His Ala Ser Glu His Ala Val Asn Lys Gln
               125                 130                 135
Leu Ala Asp Lys Glu Arg Val Ala Ala Ala Leu Glu Asn Thr His
               140                 145                 150
Leu Leu Glu Val Val Asn Gln Cys Leu Ser Ala Arg Ser
               155                 160
```

<210> 24
<211> 199
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3274451CD1

<400> 24
```
Met Thr Pro Leu Ala Pro Trp Asp Pro Lys Tyr Glu Ala Lys Ala
 1               5                  10                  15
Gly Pro Arg Pro Val Trp Gly Ala Asn Cys Ser Ser Gly Ala Ser
                20                  25                  30
Phe Ser Gly Arg Thr Leu Cys His Pro Ser Phe Trp Pro Leu Tyr
                35                  40                  45
Glu Ala Ala Ser Gly Arg Gly Leu Arg Pro Val Ala Pro Ala Thr
                50                  55                  60
Gly His Trp Asn Gly Gln Gln Ala Pro Pro Asp Ala Gly Phe Pro
                65                  70                  75
Val Val Cys Cys Glu Asp Val Phe Leu Ser Asp Pro Leu Leu Pro
                80                  85                  90
Arg Gly Gln Arg Val Pro Leu Tyr Leu Ser Lys Ala Pro Gln Gln
                95                 100                 105
Met Met Gly Ser Leu Lys Leu Leu Pro Pro Pro Pro Ile Met Ser
               110                 115                 120
Ala Arg Val Leu Pro Arg Pro Ser Pro Ser Arg Gly Pro Ser Thr
               125                 130                 135
Ala Trp Leu Ser Gly Pro Glu Leu Ile Ala Leu Thr Gly Leu Leu
               140                 145                 150
Gln Met Ser Gln Gly Glu Pro Arg Pro Ser Ser Ser Ala Val Gly
               155                 160                 165
```

Pro Pro Asp His Thr Ser Asp Pro Pro Ser Pro Cys Gly Ser Pro
170 175 180
Ser Ser Ser Gln Gly Ala Asp Leu Ser Leu Pro Gln Thr Pro Asp
185 190 195
Thr His Cys Pro

<210> 25
<211> 231
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3284256CD1

<400> 25
Met Ala Gly Met Val Asp Phe Gln Asp Glu Glu Gln Val Lys Ser
1 5 10 15
Phe Leu Glu Asn Met Glu Val Glu Cys Asn Tyr His Cys Tyr His
20 25 30
Glu Lys Asp Pro Asp Gly Cys Tyr Arg Leu Val Asp Tyr Leu Glu
35 40 45
Gly Ile Arg Lys Asn Phe Asp Glu Ala Ala Lys Val Leu Lys Phe
50 55 60
Asn Cys Glu Glu Asn Gln His Ser Asp Ser Cys Tyr Lys Leu Gly
65 70 75
Ala Tyr Tyr Val Thr Gly Lys Gly Gly Leu Thr Gln Asp Leu Lys
80 85 90
Ala Ala Ala Arg Cys Phe Leu Met Ala Cys Glu Lys Pro Gly Lys
95 100 105
Lys Ser Ile Ala Ala Cys His Asn Val Gly Leu Leu Ala His Asp
110 115 120
Gly Gln Val Asn Glu Asp Gly Gln Pro Asp Leu Gly Lys Ala Arg
125 130 135
Asp Tyr Tyr Thr Arg Ala Cys Asp Gly Gly Tyr Thr Ser Ser Cys
140 145 150
Phe Asn Leu Ser Ala Met Phe Leu Gln Gly Ala Pro Gly Phe Pro
155 160 165
Lys Asp Met Asp Leu Ala Cys Lys Tyr Ser Met Lys Ala Cys Asp
170 175 180
Leu Gly His Ile Trp Ala Cys Ala Asn Ala Ser Arg Met Tyr Lys
185 190 195
Leu Gly Asp Gly Val Asp Lys Asp Glu Ala Lys Ala Glu Val Leu
200 205 210
Lys Asn Arg Ala Gln Gln Leu His Lys Glu Gln Gln Lys Gly Val
215 220 225
Gln Pro Leu Thr Phe Gly
230

<210> 26
<211> 412
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3507004CD1

<400> 26
Met Ala Pro Val Glu His Val Val Ala Asp Ala Gly Ala Phe Leu
1 5 10 15
Arg His Ala Ala Leu Gln Asp Ile Gly Lys Asn Ile Tyr Thr Ile
20 25 30
Arg Glu Val Val Thr Glu Ile Arg Asp Lys Ala Thr Arg Arg Arg
35 40 45

WEST

Leu Ala Val Leu Pro Tyr Glu Leu Arg Phe Lys Glu Pro Leu Pro
50 55 60
Glu Tyr Val Arg Leu Val Thr Glu Phe Ser Lys Lys Thr Gly Asp
65 70 75
Tyr Pro Ser Leu Ser Ala Thr Asp Ile Gln Val Leu Ala Leu Thr
80 85 90
Tyr Gln Leu Glu Ala Glu Phe Val Gly Val Ser His Leu Lys Gln
95 100 105
Glu Pro Gln Lys Val Lys Val Ser Ser Ser Ile Gln His Pro Glu
110 115 120
Thr Pro Leu His Ile Ser Gly Phe His Leu Pro Tyr Lys Pro Lys
125 130 135
Pro Pro Gln Glu Thr Glu Lys Gly His Ser Ala Cys Glu Pro Glu
140 145 150
Asn Leu Glu Phe Ser Ser Phe Met Phe Trp Arg Asn Pro Leu Pro
155 160 165
Asn Ile Asp His Glu Leu Gln Glu Leu Leu Ile Asp Arg Gly Glu
170 175 180
Asp Val Pro Ser Glu Glu Glu Glu Glu Glu Glu Asn Gly Phe Glu
185 190 195
Asp Arg Lys Asp Asp Ser Asp Asp Asp Gly Gly Gly Trp Ile Thr
200 205 210
Pro Ser Asn Ile Lys Gln Ile Gln Gln Glu Leu Glu Gln Cys Asp
215 220 225
Val Pro Glu Asp Val Arg Val Gly Cys Leu Thr Thr Asp Phe Ala
230 235 240
Met Gln Asn Val Leu Leu Gln Met Gly Leu His Val Leu Ala Val
245 250 255
Asn Gly Met Leu Ile Arg Glu Ala Arg Ser Tyr Ile Leu Arg Cys
260 265 270
His Gly Cys Phe Lys Thr Thr Ser Asp Met Ser Arg Val Phe Cys
275 280 285
Ser His Cys Gly Asn Lys Thr Leu Lys Lys Val Ser Val Thr Val
290 295 300
Ser Asp Asp Gly Thr Leu His Met His Phe Ser Arg Asn Pro Lys
305 310 315
Val Leu Asn Pro Arg Gly Leu Arg Tyr Ser Leu Pro Thr Pro Lys
320 325 330
Gly Gly Lys Tyr Ala Ile Asn Pro His Leu Thr Glu Asp Gln Arg
335 340 345
Phe Pro Gln Leu Arg Leu Ser Gln Lys Ala Arg Gln Lys Thr Asn
350 355 360
Val Phe Ala Pro Asp Tyr Ile Ala Gly Val Ser Pro Phe Val Glu
365 370 375
Asn Asp Ile Ser Ser Arg Ser Ala Thr Leu Gln Val Arg Asp Ser
380 385 390
Thr Leu Gly Ala Gly Arg Arg Arg Leu Asn Pro Asn Ala Ser Arg
395 400 405
Lys Lys Phe Val Lys Lys Arg
410

<210> 27
<211> 272
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3585823CD1

<400> 27
Met Thr Pro Ile Leu Glu Gly Ser His Arg Ala His Ser Leu Leu
1 5 10 15
Phe Glu Asn Ser Asp Ser Phe Ser Glu Asp Ser Ser Thr Leu Gly
20 25 30
Arg Thr Arg Ser Leu Pro Ile Thr Ile Glu Met Leu Lys Val Pro

WEST

```
                    35                  40                  45
Asp Asp Glu Glu Glu Glu Glu Gln Thr Cys Pro Ser Thr Phe Ser
                50                  55                  60
Glu Glu Met Thr Pro Thr Ser Val Ile Pro Lys Leu Pro Gln Cys
                65                  70                  75
Leu Arg Glu Glu Glu Glu Lys Glu Ser Asp Ser Asp Ser Glu Gly
                80                  85                  90
Pro Ile Gln Tyr Arg Asp Glu Glu Asp Glu Asp Glu Ser Tyr Gln
                95                  100                 105
Ser Ala Leu Ala Asn Lys Val Lys Arg Lys Asp Thr Leu Ala Met
                110                 115                 120
Lys Leu Asn His Arg Pro Ser Glu Pro Glu Leu Asn Leu Asn Ser
                125                 130                 135
Trp Pro Cys Lys Ser Lys Glu Glu Trp Asn Glu Ile Arg His Gln
                140                 145                 150
Ile Gly Asn Thr Leu Ile Arg Arg Leu Ser Gln Arg Pro Thr Pro
                155                 160                 165
Glu Glu Leu Glu Gln Arg Asn Ile Leu Gln Pro Lys Asn Glu Ala
                170                 175                 180
Asp Arg Gln Ala Glu Lys Arg Glu Ile Lys Arg Arg Leu Thr Arg
                185                 190                 195
Lys Leu Ser Gln Arg Pro Thr Val Ala Glu Leu Leu Ala Arg Lys
                200                 205                 210
Ile Leu Arg Phe Asn Glu Tyr Val Glu Val Thr Asp Ala Gln Asp
                215                 220                 225
Tyr Asp Arg Arg Ala Asp Lys Pro Trp Thr Lys Leu Thr Pro Ala
                230                 235                 240
Asp Lys Ala Ala Ile Arg Lys Glu Leu Asn Glu Phe Lys Ser Ser
                245                 250                 255
Glu Met Glu Val His Glu Glu Ser Lys His Phe Thr Arg Tyr His
                260                 265                 270
Arg Pro


&lt;210&gt; 28
&lt;211&gt; 242
&lt;212&gt; PRT
&lt;213&gt; Homo sapiens

&lt;220&gt;
&lt;221&gt; misc_feature
&lt;223&gt; Incyte ID No: 3743822CD1

&lt;400&gt; 28
Met Ala Glu Ala Val Glu Arg Thr Asp Glu Leu Val Arg Glu Tyr
 1              5                   10                  15
Leu Leu Phe Arg Gly Phe Thr His Thr Leu Arg Gln Leu Asp Ala
                20                  25                  30
Glu Ile Lys Ala Asp Lys Glu Lys Gly Phe Arg Val Asp Lys Ile
                35                  40                  45
Val Asp Gln Leu Gln Gln Leu Met Gln Val Tyr Asp Leu Ala Ala
                50                  55                  60
Leu Arg Asp Tyr Trp Ser Tyr Leu Glu Arg Arg Leu Phe Ser Arg
                65                  70                  75
Leu Glu Asp Ile Tyr Arg Pro Thr Ile His Lys Leu Lys Thr Ser
                80                  85                  90
Leu Phe Arg Phe Tyr Leu Val Tyr Thr Ile Gln Thr Asn Arg Asn
                95                  100                 105
Asp Lys Ala Gln Glu Phe Phe Ala Lys Gln Ala Thr Glu Leu Gln
                110                 115                 120
Asn Gln Ala Glu Trp Lys Asp Trp Phe Val Leu Pro Phe Leu Pro
                125                 130                 135
Ser Pro Asp Thr Asn Pro Thr Phe Ala Thr Tyr Phe Ser Arg Gln
                140                 145                 150
Trp Ala Asp Thr Phe Ile Val Ser Leu His Asn Phe Leu Ser Val
                155                 160                 165
Leu Phe Gln Cys Met Pro Val Pro Val Ile Leu Asn Phe Asp Ala
```

```
                170                 175                 180
Glu Cys Gln Arg Thr Asn Gln Val Gln Glu Glu Asn Glu Val Leu
                185                 190                 195
Arg Gln Lys Leu Phe Ala Leu Gln Ala Glu Ile His Arg Leu Lys
                200                 205                 210
Lys Glu Glu Gln Gln Pro Glu Glu Glu Glu Ala Leu Val Gln His
                215                 220                 225
Lys Leu Pro Pro Tyr Val Ser Asn Met Asp Arg Leu Gly Asp Ser
                230                 235                 240
Glu Leu
```

<210> 29
<211> 285
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3808027CD1

<400> 29

```
Met Thr Thr Leu Thr His Arg Ala Arg Arg Thr Glu Ile Ser Lys
 1               5                  10                  15
Asn Ser Glu Lys Lys Met Glu Ser Glu Glu Asp Ser Asn Trp Glu
                20                  25                  30
Lys Ser Pro Asp Asn Glu Asp Ser Gly Asp Ser Lys Asp Ile Arg
                35                  40                  45
Leu Thr Leu Met Glu Glu Val Leu Leu Leu Gly Leu Lys Asp Lys
                50                  55                  60
Glu Gly Tyr Thr Ser Phe Trp Asn Asp Cys Ile Ser Ser Gly Leu
                65                  70                  75
Arg Gly Gly Ile Leu Ile Glu Leu Ala Met Arg Gly Arg Ile Tyr
                80                  85                  90
Leu Glu Pro Pro Thr Met Arg Lys Lys Arg Leu Leu Asp Arg Lys
                95                 100                 105
Val Leu Leu Lys Ser Asp Ser Pro Thr Gly Asp Val Leu Leu Asp
               110                 115                 120
Glu Thr Leu Lys His Ile Lys Ala Thr Glu Pro Thr Glu Thr Val
               125                 130                 135
Gln Thr Trp Ile Glu Leu Leu Thr Gly Glu Thr Trp Asn Pro Phe
               140                 145                 150
Lys Leu Gln Tyr Gln Leu Arg Asn Val Arg Glu Arg Ile Ala Lys
               155                 160                 165
Asn Leu Val Glu Lys Gly Ile Leu Thr Thr Glu Lys Gln Asn Phe
               170                 175                 180
Leu Leu Phe Asp Met Thr Thr His Pro Val Thr Asn Thr Thr Glu
               185                 190                 195
Lys Gln Arg Leu Val Lys Lys Leu Gln Asp Ser Val Leu Glu Arg
               200                 205                 210
Trp Val Asn Asp Pro Gln Arg Met Asp Lys Arg Thr Leu Ala Leu
               215                 220                 225
Leu Val Leu Ala His Ser Ser Asp Val Leu Glu Asn Val Phe Ser
               230                 235                 240
Ser Leu Thr Asp Asp Lys Tyr Asp Val Ala Met Asn Arg Ala Lys
               245                 250                 255
Asp Leu Val Glu Leu Asp Pro Glu Val Glu Gly Thr Lys Pro Ser
               260                 265                 270
Ala Thr Glu Met Ile Trp Ala Val Leu Ala Ala Phe Asn Lys Ser
               275                 280                 285
```

<210> 30
<211> 89
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4016220CD1

<400> 30
Met Gly Thr Lys Arg Glu Ala Ile Leu Lys Val Leu Glu Asn Leu
1 5 10 15
Thr Pro Glu Glu Leu Lys Lys Phe Lys Met Lys Leu Gly Thr Val
20 25 30
Pro Leu Arg Glu Gly Phe Glu Arg Ile Pro Arg Gly Ala Leu Gly
35 40 45
Gln Leu Asp Ile Val Asp Leu Thr Asp Lys Leu Val Ala Ser Tyr
50 55 60
Tyr Glu Asp Tyr Ala Ala Glu Leu Val Val Ala Val Leu Arg Asp
65 70 75
Met Arg Met Leu Glu Glu Ala Ala Arg Leu Gln Arg Ala Ala
80 85

<210> 31
<211> 210
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4093555CD1

<400> 31
Met Ala Thr Glu Thr Val Glu Leu His Lys Leu Lys Leu Ala Glu
1 5 10 15
Leu Lys Gln Glu Cys Leu Ala Arg Gly Leu Glu Thr Lys Gly Ile
20 25 30
Lys Gln Asp Leu Ile His Arg Leu Gln Ala Tyr Leu Glu Glu His
35 40 45
Ala Glu Glu Glu Ala Asn Glu Glu Asp Val Leu Gly Asp Glu Thr
50 55 60
Glu Glu Glu Glu Thr Lys Pro Ile Glu Leu Pro Val Lys Glu Glu
65 70 75
Glu Pro Pro Glu Lys Thr Val Asp Val Ala Ala Glu Lys Lys Val
80 85 90
Val Lys Ile Thr Ser Glu Ile Pro Gln Thr Glu Arg Met Gln Lys
95 100 105
Arg Ala Glu Arg Phe Asn Val Pro Val Ser Leu Glu Ser Lys Lys
110 115 120
Ala Ala Arg Ala Ala Arg Phe Gly Ile Ser Ser Val Pro Thr Lys
125 130 135
Gly Leu Ser Ser Asp Asn Lys Pro Met Val Asn Leu Asp Lys Leu
140 145 150
Lys Glu Arg Ala Gln Arg Phe Gly Leu Asn Val Ser Ser Ile Ser
155 160 165
Arg Lys Ser Glu Asp Asp Glu Lys Leu Lys Lys Arg Lys Glu Arg
170 175 180
Phe Gly Ile Val Thr Ser Ser Ala Gly Thr Gly Thr Thr Glu Asp
185 190 195
Thr Glu Ala Lys Lys Arg Lys Arg Ala Glu Arg Phe Gly Ile Ala
200 205 210

<210> 32
<211> 271
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4829366CD1

<400> 32
Met Ala Ala Glu Glu Pro Gln Gln Gln Lys Gln Glu Pro Leu Gly
1 5 10 15
Ser Asp Ser Glu Gly Val Asn Cys Leu Ala Tyr Asp Glu Ala Ile
20 25 30
Met Ala Gln Gln Asp Arg Ile Gln Gln Glu Ile Ala Val Gln Asn
35 40 45
Pro Leu Val Ser Glu Arg Leu Glu Leu Ser Val Leu Tyr Lys Glu
50 55 60
Tyr Ala Glu Asp Asp Asn Ile Tyr Gln Gln Lys Ile Lys Asp Leu
65 70 75
His Lys Lys Tyr Ser Tyr Ile Arg Lys Thr Arg Pro Asp Gly Asn
80 85 90
Cys Phe Tyr Arg Ala Phe Gly Phe Ser His Leu Glu Ala Leu Leu
95 100 105
Asp Asp Ser Lys Glu Leu Gln Arg Phe Lys Ala Val Ser Ala Lys
110 115 120
Ser Lys Glu Asp Leu Val Ser Gln Gly Phe Thr Glu Phe Thr Ile
125 130 135
Glu Asp Phe His Asn Thr Phe Met Asp Leu Ile Glu Gln Val Glu
140 145 150
Lys Gln Thr Ser Val Ala Asp Leu Leu Ala Ser Phe Asn Asp Gln
155 160 165
Ser Thr Ser Asp Tyr Leu Val Val Tyr Leu Arg Leu Leu Thr Ser
170 175 180
Gly Tyr Leu Gln Arg Glu Ser Lys Phe Phe Glu His Phe Ile Glu
185 190 195
Gly Gly Arg Thr Val Lys Glu Phe Cys Gln Gln Glu Val Glu Pro
200 205 210
Met Cys Lys Glu Ser Asp His Ile His Ile Ile Ala Leu Ala Gln
215 220 225
Ala Leu Ser Val Ser Ile Gln Val Glu Tyr Met Asp Arg Gly Glu
230 235 240
Gly Gly Thr Thr Asn Pro His Ile Phe Pro Glu Gly Ser Glu Pro
245 250 255
Lys Val Tyr Leu Leu Tyr Arg Pro Gly His Tyr Asp Ile Leu Tyr
260 265 270
Lys

<210> 33
<211> 389
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4830720CD1

<400> 33
Met Glu Ala Leu Gly Lys Leu Lys Gln Phe Asp Ala Tyr Pro Lys
1 5 10 15
Thr Leu Glu Asp Phe Arg Val Lys Thr Cys Gly Gly Ala Thr Val
20 25 30
Thr Ile Val Ser Gly Leu Leu Met Leu Leu Leu Phe Leu Ser Glu
35 40 45
Leu Gln Tyr Tyr Leu Thr Thr Glu Val His Pro Glu Leu Tyr Val
50 55 60
Asp Lys Ser Arg Gly Asp Lys Leu Lys Ile Asn Ile Asp Val Leu
65 70 75
Phe Pro His Met Pro Cys Ala Tyr Leu Ser Ile Asp Ala Met Asp
80 85 90
Val Ala Gly Glu Gln Gln Leu Asp Val Glu His Asn Leu Leu Lys
95 100 105
Gln Arg Leu Asp Lys Asp Gly Ile Pro Val Ser Ser Glu Ala Glu
110 115 120
Arg His Glu Leu Gly Lys Val Glu Val Thr Val Phe Asp Pro Asp

```
                125                     130                     135
Ser Leu Asp Pro Asp Arg Cys Glu Ser Cys Tyr Gly Ala Glu Ala
                140                     145                     150
Glu Asp Ile Lys Cys Cys Asn Thr Cys Glu Asp Val Arg Glu Ala
                155                     160                     165
Tyr Arg Arg Arg Gly Trp Ala Phe Lys Asn Pro Asp Thr Ile Glu
                170                     175                     180
Gln Cys Arg Arg Glu Gly Phe Ser Gln Lys Met Gln Glu Gln Lys
                185                     190                     195
Asn Glu Gly Cys Gln Val Tyr Gly Phe Leu Glu Val Asn Lys Val
                200                     205                     210
Ala Gly Asn Phe His Phe Ala Pro Gly Lys Ser Phe Gln Gln Ser
                215                     220                     225
His Val His Val His Asp Leu Gln Ser Phe Gly Leu Asp Asn Ile
                230                     235                     240
Asn Met Thr His Tyr Ile Gln His Leu Ser Phe Gly Glu Asp Tyr
                245                     250                     255
Pro Gly Ile Val Asn Pro Leu Asp His Thr Asn Val Thr Ala Pro
                260                     265                     270
Gln Ala Ser Met Met Phe Gln Tyr Phe Val Lys Val Val Pro Thr
                275                     280                     285
Val Tyr Met Lys Val Asp Gly Glu Ala Pro Leu Pro Pro Gln Val
                290                     295                     300
Leu Arg Thr Asn Gln Phe Ser Val Thr Arg His Glu Lys Val Ala
                305                     310                     315
Asn Gly Leu Leu Gly Asp Gln Gly Leu Pro Gly Val Phe Val Leu
                320                     325                     330
Tyr Glu Leu Ser Pro Met Met Val Lys Leu Thr Glu Lys His Arg
                335                     340                     345
Ser Phe Thr His Phe Leu Thr Gly Val Cys Ala Ile Ile Gly Gly
                350                     355                     360
Met Phe Thr Val Ala Gly Leu Ile Asp Ser Leu Ile Tyr His Ser
                365                     370                     375
Ala Arg Ala Ile Gln Lys Lys Ile Asp Leu Gly Lys Thr Thr
                380                     385
```

```
<210> 34
<211> 228
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 5389730CD1

<400> 34
Met Ser Tyr Ala Glu Lys Pro Asp Glu Ile Thr Lys Asp Glu Trp
 1              5                       10                      15
Met Glu Lys Leu Asn Asn Leu His Val Gln Arg Ala Asp Met Asn
                20                      25                      30
Arg Leu Ile Met Asn Tyr Leu Val Thr Glu Gly Phe Lys Glu Ala
                35                      40                      45
Ala Glu Lys Phe Arg Met Glu Ser Gly Ile Glu Pro Ser Val Asp
                50                      55                      60
Leu Glu Thr Leu Asp Glu Arg Ile Lys Ile Arg Glu Met Ile Leu
                65                      70                      75
Lys Gly Gln Ile Gln Glu Ala Ile Ala Leu Ile Asn Ser Leu His
                80                      85                      90
Pro Glu Leu Leu Asp Thr Asn Arg Tyr Leu Tyr Phe His Leu Gln
                95                      100                     105
Gln Gln His Leu Ile Glu Leu Ile Arg Gln Arg Glu Thr Glu Ala
                110                     115                     120
Ala Leu Glu Phe Ala Gln Thr Gln Leu Ala Glu Gln Gly Glu Glu
                125                     130                     135
Ser Arg Glu Cys Leu Thr Glu Met Glu Arg Thr Leu Ala Leu Leu
                140                     145                     150
```


WEST

```
Ala Phe Asp Ser Pro Glu Glu Ser Pro Phe Gly Asp Leu Leu His
                155                 160                 165
Thr Met Gln Arg Gln Lys Val Trp Ser Glu Val Asn Gln Ala Val
                170                 175                 180
Leu Asp Tyr Glu Asn Arg Glu Ser Thr Pro Lys Leu Ala Lys Leu
                185                 190                 195
Leu Lys Leu Leu Leu Trp Ala Gln Asn Glu Leu Asp Gln Lys Lys
                200                 205                 210
Val Lys Tyr Pro Lys Met Thr Asp Leu Ser Lys Gly Val Ile Glu
                215                 220                 225
Glu Pro Lys
```

<210> 35
<211> 330
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 5397088CD1

<400> 35

```
Met Ala Pro Glu Glu Asn Ala Gly Thr Glu Leu Leu Leu Gln Ser
 1              5                   10                  15
Phe Glu Arg Arg Phe Leu Ala Ala Arg Thr Leu Arg Ser Phe Pro
                20                  25                  30
Trp Gln Ser Leu Glu Ala Lys Leu Arg Asp Ser Ser Asp Ser Glu
                35                  40                  45
Leu Leu Arg Asp Ile Leu His Lys Thr Val Lys His Pro Val Cys
                50                  55                  60
Val Lys His Pro Pro Ser Val Lys Tyr Ala Arg Cys Phe Leu Ser
                65                  70                  75
Glu Leu Ile Lys Lys His Glu Ala Val His Thr Glu Pro Leu Asp
                80                  85                  90
Glu Leu Tyr Glu Ala Leu Ala Glu Thr Leu Met Ala Lys Glu Ser
                95                  100                 105
Thr Gln Gly His Arg Ser Tyr Leu Leu Pro Ser Gly Gly Ser Val
                110                 115                 120
Thr Leu Ser Glu Ser Thr Ala Ile Ile Ser Tyr Gly Thr Thr Gly
                125                 130                 135
Leu Val Thr Trp Asp Ala Ala Leu Tyr Leu Ala Glu Trp Ala Ile
                140                 145                 150
Glu Asn Pro Ala Val Phe Thr Asn Arg Thr Val Leu Glu Leu Gly
                155                 160                 165
Ser Gly Ala Gly Leu Thr Gly Leu Ala Ile Cys Lys Met Cys Arg
                170                 175                 180
Pro Arg Ala Tyr Ile Phe Ser Asp Cys His Ser Arg Val Leu Glu
                185                 190                 195
Gln Leu Arg Gly Asn Val Leu Leu Asn Gly Leu Ser Leu Glu Ala
                200                 205                 210
Asp Ile Thr Ala Lys Leu Asp Ser Pro Arg Val Thr Val Ala Gln
                215                 220                 225
Leu Asp Trp Asp Val Ala Thr Val His Gln Leu Ser Ala Phe Gln
                230                 235                 240
Pro Asp Val Val Ile Ala Ala Asp Val Leu Tyr Cys Pro Glu Ala
                245                 250                 255
Ile Met Ser Leu Val Gly Val Leu Arg Arg Leu Ala Ala Cys Arg
                260                 265                 270
Glu Asp Gln Arg Ala Pro Glu Val Tyr Val Ala Phe Thr Val Arg
                275                 280                 285
Asn Pro Glu Thr Cys Gln Leu Phe Thr Thr Glu Leu Gly Arg Ala
                290                 295                 300
Gly Ile Arg Trp Glu Val Glu Pro Arg His Glu Gln Lys Leu Phe
                305                 310                 315
Pro Tyr Glu Glu His Leu Glu Met Ala Met Leu Asn Leu Thr Leu
                320                 325                 330
```

<210> 36
<211> 307
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 5425521CD1

<400> 36
Met Ala Ala Leu Ala Pro Leu Pro Pro Leu Pro Ala Gln Phe Lys
1 5 10 15
Ser Ile Gln His His Leu Arg Thr Ala Gln Glu His Asp Lys Arg
20 25 30
Asp Pro Val Val Ala Tyr Tyr Cys Arg Leu Tyr Ala Met Gln Thr
35 40 45
Gly Met Lys Ile Asp Ser Lys Thr Pro Glu Cys Arg Lys Phe Leu
50 55 60
Ser Lys Leu Met Asp Gln Leu Glu Ala Leu Lys Lys Gln Leu Gly
65 70 75
Asp Asn Glu Ala Ile Thr Gln Glu Ile Val Gly Cys Ala His Leu
80 85 90
Glu Asn Tyr Ala Leu Lys Met Phe Leu Tyr Ala Asp Asn Glu Asp
95 100 105
Arg Ala Gly Arg Phe His Lys Asn Met Ile Lys Ser Phe Tyr Thr
110 115 120
Ala Ser Leu Leu Ile Asp Val Ile Thr Val Phe Gly Glu Leu Thr
125 130 135
Asp Glu Asn Val Lys His Arg Lys Tyr Ala Arg Trp Lys Ala Thr
140 145 150
Tyr Ile His Asn Cys Leu Lys Asn Gly Glu Thr Pro Gln Ala Gly
155 160 165
Pro Val Gly Ile Glu Glu Asp Asn Asp Ile Glu Glu Asn Glu Asp
170 175 180
Ala Gly Ala Ala Ser Leu Pro Thr Gln Pro Thr Gln Pro Ser Ser
185 190 195
Ser Ser Thr Tyr Asp Pro Ser Asn Met Pro Ser Gly Asn Tyr Thr
200 205 210
Gly Ile Gln Ile Pro Pro Gly Ala His Ala Pro Ala Asn Thr Pro
215 220 225
Ala Glu Val Pro His Ser Thr Gly Val Ala Ser Asn Thr Ile Gln
230 235 240
Pro Thr Pro Gln Thr Ile Pro Ala Ile Asp Pro Ala Leu Phe Asn
245 250 255
Thr Ile Ser Gln Gly Asp Val Arg Leu Thr Pro Glu Asp Phe Ala
260 265 270
Arg Ala Gln Lys Tyr Cys Lys Tyr Ala Gly Ser Ala Leu Gln Tyr
275 280 285
Glu Asp Val Ser Thr Ala Val Gln Asn Leu Gln Lys Ala Leu Lys
290 295 300
Leu Leu Thr Thr Gly Arg Glu
305

<210> 37
<211> 163
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 5495427CD1

<400> 37
Met Ala Trp Ala Val Ser His Phe Arg Pro Gly Pro Glu Val Trp
1 5 10 15

```
Asp Thr Ala Ser Met Ala Ala Ser Lys Val Lys Gln Asp Met Pro
                20                  25                  30
Pro Pro Gly Gly Tyr Gly Pro Ile Asp Tyr Lys Arg Asn Leu Pro
                35                  40                  45
Arg Arg Gly Leu Ser Gly Tyr Ser Met Leu Ala Ile Gly Ile Gly
                50                  55                  60
Thr Leu Ile Tyr Gly His Trp Ser Ile Met Lys Trp Asn Arg Glu
                65                  70                  75
Arg Arg Arg Leu Gln Ile Glu Asp Phe Glu Ala Arg Ile Ala Leu
                80                  85                  90
Leu Pro Leu Leu Gln Ala Glu Thr Asp Arg Arg Thr Leu Gln Met
                95                  100                 105
Leu Arg Glu Asn Leu Glu Glu Glu Ala Ile Ile Met Lys Asp Val
                110                 115                 120
Pro Asp Trp Lys Val Gly Glu Ser Val Phe His Thr Thr Arg Trp
                125                 130                 135
Val Pro Pro Leu Ile Gly Glu Leu Tyr Gly Leu Arg Thr Thr Glu
                140                 145                 150
Glu Ala Leu His Ala Ser His Gly Phe Met Trp Tyr Thr
                155                 160
```

```
<210> 38
<211> 202
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 003908CD1

<400> 38
Met Ala Val Pro Ala Ala Leu Ile Leu Arg Glu Ser Pro Ser Met
 1               5                  10                  15
Lys Lys Ala Val Ser Leu Ile Asn Ala Ile Asp Thr Gly Arg Phe
                20                  25                  30
Pro Arg Leu Leu Thr Arg Ile Leu Gln Lys Leu His Leu Lys Ala
                35                  40                  45
Glu Ser Ser Phe Ser Glu Glu Glu Glu Glu Lys Leu Gln Ala Ala
                50                  55                  60
Phe Ser Leu Glu Lys Gln Asp Leu His Leu Val Leu Glu Thr Ile
                65                  70                  75
Ser Phe Ile Leu Glu Gln Ala Val Tyr His Asn Val Lys Pro Ala
                80                  85                  90
Ala Leu Gln Gln Gln Leu Glu Asn Ile His Leu Arg Gln Asp Lys
                95                  100                 105
Ala Glu Ala Phe Val Asn Thr Trp Ser Ser Met Gly Gln Glu Thr
                110                 115                 120
Val Glu Lys Phe Arg Gln Arg Ile Leu Ala Pro Cys Lys Leu Glu
                125                 130                 135
Thr Val Gly Trp Gln Leu Asn Leu Gln Met Ala His Ser Ala Gln
                140                 145                 150
Ala Lys Leu Lys Ser Pro Gln Ala Val Leu Gln Leu Gly Val Asn
                155                 160                 165
Asn Glu Asp Ser Lys Ser Leu Glu Lys Val Leu Val Glu Phe Ser
                170                 175                 180
His Lys Glu Leu Phe Asp Phe Tyr Asn Lys Leu Glu Thr Ile Gln
                185                 190                 195
Ala Gln Leu Asp Ser Leu Thr
                200
```

```
<210> 39
<211> 209
<212> PRT
<213> Homo sapiens
```

<220>
<221> misc_feature
<223> Incyte ID No: 085915CD1

<400> 39
Met Ile Asp Asn Phe Phe Ser Glu Pro Thr Thr Lys Ser Trp Glu
1 5 10 15
Ile Ile Thr Val Glu Glu Ala Lys Arg Arg Lys Ser Thr Cys Ser
20 25 30
Tyr Tyr Glu Asp Glu Asp Glu Glu Val Leu Pro Val Leu Arg Pro
35 40 45
His Ser Ala Leu Leu Glu Asn Met His Ile Glu Gln Leu Ala Arg
50 55 60
Arg Leu Pro Ala Arg Val Gln Gly Tyr Pro Trp Arg Leu Ala Tyr
65 70 75
Ser Thr Leu Glu His Gly Thr Ser Leu Lys Thr Leu Tyr Arg Lys
80 85 90
Ser Ala Ser Leu Asp Ser Pro Val Leu Leu Val Ile Lys Asp Met
95 100 105
Asp Asn Gln Ile Phe Gly Ala Tyr Ala Thr His Pro Phe Lys Phe
110 115 120
Ser Asp His Tyr Tyr Gly Thr Gly Glu Thr Phe Leu Tyr Thr Phe
125 130 135
Ser Pro His Phe Lys Val Phe Lys Trp Ser Gly Glu Asn Ser Tyr
140 145 150
Phe Ile Asn Gly Asp Ile Ser Ser Leu Glu Leu Gly Gly Gly Gly
155 160 165
Gly Arg Phe Gly Leu Trp Leu Asp Ala Asp Leu Tyr His Gly Arg
170 175 180
Ser Asn Ser Cys Ser Thr Phe Asn Asn Asp Ile Leu Ser Lys Lys
185 190 195
Glu Asp Phe Ile Val Gln Asp Leu Glu Val Trp Ala Phe Asp
200 205

<210> 40
<211> 314
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 478615CD1

<400> 40
Met Pro Gln Met Arg Gln Thr Pro Thr Asp Lys Pro Leu Cys Pro
1 5 10 15
Ser Arg Thr His Lys Val Leu Pro Ile Leu Glu Ile Leu Tyr His
20 25 30
Val Glu Glu Arg Asn Ser His His Val Tyr Met Ala Leu Ile Ile
35 40 45
Leu Leu Ile Leu Thr Glu Asp Asp Gly Phe Asn Arg Ser Ile His
50 55 60
Glu Val Ile Leu Lys Asn Ile Thr Trp Tyr Ser Glu Arg Val Leu
65 70 75
Thr Glu Ile Ser Leu Gly Ser Leu Leu Ile Leu Val Val Ile Arg
80 85 90
Thr Ile Gln Tyr Asn Met Thr Arg Thr Arg Asp Lys Tyr Leu His
95 100 105
Thr Asn Cys Leu Ala Ala Leu Ala Asn Met Ser Ala Gln Phe Arg
110 115 120
Ser Leu His Gln Tyr Ala Ala Gln Arg Ile Ile Ser Leu Phe Ser
125 130 135
Leu Leu Ser Lys Lys His Asn Lys Val Leu Glu Gln Ala Thr Gln
140 145 150
Ser Leu Arg Gly Ser Leu Ser Ser Asn Asp Val Pro Leu Pro Asp
155 160 165

```
Tyr Ala Gln Asp Leu Asn Val Ile Glu Glu Val Ile Arg Met Met
                170                 175                 180
Leu Glu Ile Ile Asn Ser Cys Leu Thr Asn Ser Leu His His Asn
                185                 190                 195
Pro Asn Leu Val Tyr Ala Leu Leu Tyr Lys Arg Asp Leu Phe Glu
                200                 205                 210
Gln Phe Arg Thr His Pro Ser Phe Gln Asp Ile Met Gln Asn Ile
                215                 220                 225
Asp Leu Val Ile Ser Phe Phe Ser Ser Arg Leu Leu Gln Ala Gly
                230                 235                 240
Ala Glu Leu Ser Val Glu Arg Val Leu Glu Ile Ile Lys Gln Gly
                245                 250                 255
Val Val Ala Leu Pro Lys Asp Arg Leu Lys Lys Phe Pro Glu Leu
                260                 265                 270
Lys Phe Lys Tyr Val Glu Glu Glu Gln Pro Glu Glu Phe Phe Ile
                275                 280                 285
Pro Tyr Val Trp Ser Leu Val Tyr Asn Ser Ala Val Gly Leu Tyr
                290                 295                 300
Trp Asn Pro Gln Asp Ile Gln Leu Phe Thr Met Asp Ser Asp
                305                 310


<210> 41
<211> 137
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 924880CD1

<400> 41
Met Glu Val Leu Glu Met Arg Ala Lys Asn Pro Val Pro Gln Leu
 1               5                  10                  15
Arg Lys Phe Lys Thr Asn Val Leu Pro Phe Arg Gln Asn Asp Ser
                20                  25                  30
Ser Ser His Cys Gln Lys Ser Gly Ser Pro Ile Ser Ser Glu Glu
                35                  40                  45
Arg Arg Arg Arg Asp Lys Gln His Leu Asp Asp Ile Thr Ala Ala
                50                  55                  60
Arg Leu Leu Pro Leu His His Met Pro Thr Gln Leu Leu Ser Ile
                65                  70                  75
Glu Glu Ser Leu Ala Leu Gln Lys Gln Gln Lys Gln Asn Tyr Glu
                80                  85                  90
Glu Met Gln Ala Lys Leu Ala Ala Gln Lys Leu Ala Glu Arg Leu
                95                  100                 105
Asn Ile Lys Met Arg Ser Tyr Asn Pro Glu Gly Glu Ser Ser Gly
                110                 115                 120
Arg Tyr Arg Glu Val Arg Asp Glu Asp Asp Asp Trp Ser Ser Asp
                125                 130                 135
Glu Phe


<210> 42
<211> 245
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 955431CD1

<400> 42
Met Ser Asp Glu Phe Ser Leu Ala Asp Ala Leu Pro Glu His Ser
 1               5                  10                  15
Pro Ala Lys Thr Ser Ala Val Ser Asn Thr Lys Pro Gly Gln Pro
                20                  25                  30
```

```
Pro Gln Gly Trp Pro Gly Ser Asn Pro Trp Asn Asn Pro Ser Ala
                35                  40                  45
Pro Ser Ser Val Pro Ser Gly Leu Pro Pro Ser Ala Thr Pro Ser
                50                  55                  60
Thr Val Pro Phe Gly Pro Ala Pro Thr Gly Met Tyr Pro Ser Val
                65                  70                  75
Pro Pro Thr Gly Pro Pro Pro Gly Pro Pro Ala Pro Phe Pro Pro
                80                  85                  90
Ser Gly Pro Ser Cys Pro Pro Pro Gly Gly Pro Tyr Pro Ala Pro
                95                  100                 105
Thr Val Pro Gly Pro Gly Pro Thr Gly Pro Tyr Pro Thr Pro Asn
                110                 115                 120
Met Pro Phe Pro Glu Leu Pro Arg Pro Tyr Gly Ala Pro Thr Asp
                125                 130                 135
Pro Ala Ala Ala Gly Pro Leu Gly Pro Trp Gly Ser Met Ser Ser
                140                 145                 150
Gly Pro Trp Ala Pro Gly Met Gly Gly Gln Tyr Pro Thr Pro Asn
                155                 160                 165
Met Pro Tyr Pro Ser Pro Gly Pro Tyr Pro Ala Pro Pro Pro Pro
                170                 175                 180
Gln Ala Pro Gly Ala Ala Pro Pro Val Pro Trp Gly Thr Val Pro
                185                 190                 195
Pro Gly Ala Trp Gly Pro Pro Ala Pro Tyr Pro Ala Pro Thr Gly
                200                 205                 210
Ser Tyr Pro Thr Pro Gly Leu Tyr Pro Thr Pro Ser Asn Pro Phe
                215                 220                 225
Gln Val Pro Ser Gly Pro Ser Gly Ala Pro Pro Met Pro Gly Gly
                230                 235                 240
Pro His Ser Tyr His
                245
```

<210> 43
<211> 179
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1275918CD1

<400> 43

```
Met Glu Thr Lys Asp Gln Lys Lys Gln Arg Lys Lys Asn Ser Gly
 1              5                   10                  15
Pro Lys Ala Ala Lys Lys Lys Lys Arg His Leu Gln Asp Leu Gln
                20                  25                  30
Leu Gly Asp Glu Glu Asp Ala Trp Lys Arg Asn Pro Lys Ala Phe
                35                  40                  45
Ala Phe Gln Ser Ala Val Trp Met Ala Arg Ser Phe His Arg Thr
                50                  55                  60
Gln Asp Leu Lys Thr Lys Lys His His Ile Pro Val Val Asp Arg
                65                  70                  75
Thr Pro Leu Glu Pro Pro Pro Ile Val Val Val Val Met Gly Pro
                80                  85                  90
Pro Lys Val Gly Lys Ser Thr Leu Ile Gln Cys Leu Ile Arg Asn
                95                  100                 105
Phe Thr Arg Gln Lys Leu Thr Lys Ile Arg Gly Pro Val Met Ile
                110                 115                 120
Val Ser Gly Lys Lys Leu Arg Leu Thr Ile Ile Glu Cys Gly Cys
                125                 130                 135
Asp Ile Asn Met Met Ile Asp Leu Ala Glu Val Ala Asp Leu Val
                140                 145                 150
Ser Glu Gln Gly Gln Pro Gly Val Leu Met Glu Thr Tyr Ser Ile
                155                 160                 165
Val Ile Gly Tyr Leu Pro Arg Asp Glu Gly Asn Arg Val Leu
                170                 175
```

<210> 44
<211> 165
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1290896CD1

<400> 44

```
Met Leu Ser Leu Cys Ala Gln Asp Leu Thr Gln Met Leu Ala Leu
 1               5                  10                  15
Ser Arg His Ser Leu Leu Ser Pro Leu Leu Ser Val Thr Ser Phe
                20                  25                  30
Arg Arg Phe Tyr Arg Gly Asp Ser Pro Thr Asp Ser Gln Lys Asp
                35                  40                  45
Met Ile Glu Ile Pro Leu Pro Pro Trp Gln Glu Arg Thr Asp Glu
                50                  55                  60
Ser Ile Glu Thr Lys Arg Ala Arg Leu Leu Tyr Glu Ser Arg Lys
                65                  70                  75
Arg Gly Met Leu Glu Asn Cys Ile Leu Leu Ser Leu Phe Ala Lys
                80                  85                  90
Glu His Leu Gln His Met Thr Glu Lys Gln Leu Asn Leu Tyr Asp
                95                 100                 105
Arg Leu Ile Asn Glu Pro Ser Asn Asp Trp Asp Ile Tyr Tyr Trp
               110                 115                 120
Ala Thr Glu Ala Lys Pro Ala Pro Glu Ile Phe Glu Asn Glu Val
               125                 130                 135
Met Ala Leu Leu Arg Asp Phe Ala Lys Asn Lys Asn Lys Glu Gln
               140                 145                 150
Arg Leu Arg Ala Pro Asp Leu Glu Tyr Leu Phe Glu Lys Pro Arg
               155                 160                 165
```

<210> 45
<211> 177
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1342736CD1

<400> 45

```
Met Gly Val Ser Val Asp Val His Gln Val Tyr Lys Tyr Pro Phe
 1               5                  10                  15
Glu Gln Val Val Ala Ser Phe Leu Arg Lys Tyr Pro Asn Pro Met
                20                  25                  30
Asp Lys Asn Val Ile Ser Val Lys Ile Met Glu Glu Lys Arg Asp
                35                  40                  45
Glu Ser Thr Gly Val Ile Tyr Arg Lys Arg Ile Ala Ile Cys Gln
                50                  55                  60
Asn Val Val Pro Glu Ile Leu Arg Lys Val Ser Ile Leu Lys Val
                65                  70                  75
Pro Asn Ile Gln Leu Glu Glu Glu Ser Trp Leu Asn Pro Arg Glu
                80                  85                  90
Arg Asn Met Ala Ile Arg Ser His Cys Leu Thr Trp Thr Gln Tyr
                95                 100                 105
Ala Ser Met Lys Glu Glu Ser Val Phe Arg Glu Ser Met Glu Asn
               110                 115                 120
Pro Asn Trp Thr Glu Phe Ile Gln Arg Gly Arg Ile Ser Ile Thr
               125                 130                 135
Gly Val Gly Phe Leu Asn Cys Val Leu Glu Thr Phe Ala Ser Thr
               140                 145                 150
Phe Leu Arg Gln Gly Ala Gln Lys Gly Ile Arg Ile Met Glu Met
               155                 160                 165
Leu Leu Lys Glu Gln Cys Gly Ala Pro Leu Ala Glu
```

170 175

<210> 46
<211> 215
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1394209CD1

<400> 46
Met Ser Ile Leu Glu Thr Ile Thr Ser Leu Asn Gln Glu Ala Ser
1 5 10 15
Ala Ala Arg Ala Ser Thr Glu Thr Ser Asn Ala Lys Thr Ser Glu
20 25 30
Arg Ala Ser Lys Lys Leu Pro Ser Gln Pro Thr Thr Asp Thr Ser
35 40 45
Thr Asp Lys Glu Arg Thr Ser Glu Asp Met Ala Asp Lys Glu Lys
50 55 60
Ser Thr Ala Asp Ser Gly Gly Glu Gly Leu Glu Thr Ala Pro Lys
65 70 75
Ser Glu Glu Phe Ser Asp Leu Pro Cys Pro Val Glu Glu Ile Lys
80 85 90
Asn Tyr Thr Lys Glu His Asn Asn Leu Ile Leu Leu Asn Lys Asp
95 100 105
Val Gln Gln Glu Ser Ser Glu Gln Lys Asn Lys Ser Thr Asp Lys
110 115 120
Gly Glu Lys Lys Pro Asp Ser Asn Glu Lys Gly Glu Arg Lys Lys
125 130 135
Glu Lys Lys Glu Lys Thr Glu Lys Lys Phe Asp His Ser Lys Lys
140 145 150
Ser Glu Asp Thr Gln Lys Val Lys Asp Glu Lys Gln Ala Lys Glu
155 160 165
Lys Glu Val Glu Ser Leu Lys Leu Pro Ser Glu Lys Asn Ser Asn
170 175 180
Lys Ala Lys Thr Val Glu Gly Thr Lys Glu Glu Glu Asn Lys Thr
185 190 195
Gln Lys Lys Lys Ser Ser Tyr Tyr Lys Asn Ile Leu Arg Ala Gln
200 205 210
Leu Leu Asn Tyr Thr
215

<210> 47
<211> 133
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1394647CD1

<400> 47
Met Leu Ser Pro Glu Ala Glu Arg Val Leu Arg Tyr Leu Val Glu
1 5 10 15
Val Glu Glu Leu Ala Glu Glu Val Leu Ala Asp Lys Arg Gln Ile
20 25 30
Val Asp Leu Asp Thr Lys Arg Asn Gln Asn Arg Glu Gly Leu Arg
35 40 45
Ala Leu Gln Lys Asp Leu Ser Leu Ser Glu Asp Val Met Val Cys
50 55 60
Phe Gly Asn Met Phe Ile Lys Met Pro His Pro Glu Thr Lys Glu
65 70 75
Met Ile Glu Lys Asp Gln Asp His Leu Asp Lys Glu Ile Glu Lys
80 85 90

```
Leu Arg Lys Gln Leu Lys Val Lys Val Asn Arg Leu Phe Glu Ala
                95                  100                 105
Gln Gly Lys Pro Glu Leu Lys Gly Phe Asn Leu Asn Pro Leu Asn
                110                 115                 120
Gln Asp Glu Leu Lys Ala Leu Lys Val Ile Leu Lys Gly
                125                 130
```

<210> 48
<211> 579
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1436854CD1

<400> 48

```
Met Ser Ala Ser Ser Leu Leu Glu Gln Arg Pro Lys Gly Gln Gly
 1              5                   10                  15
Asn Lys Val Gln Asn Gly Ser Val His Gln Lys Asp Gly Leu Asn
                20                  25                  30
Asp Asp Asp Phe Glu Pro Tyr Leu Ser Pro Gln Ala Arg Pro Asn
                35                  40                  45
Asn Ala Tyr Thr Ala Met Ser Asp Ser Tyr Leu Pro Ser Tyr Tyr
                50                  55                  60
Ser Pro Ser Ile Gly Phe Ser Tyr Ser Leu Gly Glu Ala Ala Trp
                65                  70                  75
Ser Thr Gly Gly Asp Thr Ala Met Pro Tyr Leu Thr Ser Tyr Gly
                80                  85                  90
Gln Leu Ser Asn Gly Glu Pro His Phe Leu Pro Asp Ala Met Phe
                95                  100                 105
Gly Gln Pro Gly Ala Leu Gly Ser Thr Pro Phe Leu Gly Gln His
                110                 115                 120
Gly Phe Asn Phe Phe Pro Ser Gly Ile Asp Phe Ser Ala Trp Gly
                125                 130                 135
Asn Asn Ser Ser Gln Gly Gln Ser Thr Gln Ser Ser Gly Tyr Ser
                140                 145                 150
Ser Asn Tyr Ala Tyr Ala Pro Ser Ser Leu Gly Gly Ala Met Ile
                155                 160                 165
Asp Gly Gln Ser Ala Phe Ala Asn Glu Thr Leu Asn Lys Ala Pro
                170                 175                 180
Gly Met Asn Thr Ile Asp Gln Gly Met Ala Ala Leu Lys Leu Gly
                185                 190                 195
Ser Thr Glu Val Ala Ser Asn Val Pro Lys Val Val Gly Ser Ala
                200                 205                 210
Val Gly Ser Gly Ser Ile Thr Ser Asn Ile Val Ala Ser Asn Ser
                215                 220                 225
Leu Pro Pro Ala Thr Ile Ala Pro Pro Lys Pro Ala Ser Trp Ala
                230                 235                 240
Asp Ile Ala Ser Lys Pro Ala Lys Gln Gln Pro Lys Leu Lys Thr
                245                 250                 255
Lys Asn Gly Ile Ala Gly Ser Ser Leu Pro Pro Pro Pro Ile Lys
                260                 265                 270
His Asn Met Asp Ile Gly Thr Trp Asp Asn Lys Gly Pro Val Ala
                275                 280                 285
Lys Ala Pro Ser Gln Ala Leu Val Gln Asn Ile Gly Gln Pro Thr
                290                 295                 300
Gln Gly Ser Pro Gln Pro Val Gly Gln Gln Ala Asn Asn Ser Pro
                305                 310                 315
Pro Val Ala Gln Ala Ser Val Gly Gln Gln Thr Gln Pro Leu Pro
                320                 325                 330
Pro Pro Pro Pro Gln Pro Ala Gln Leu Ser Val Gln Gln Gln Ala
                335                 340                 345
Ala Gln Pro Thr Arg Trp Val Ala Pro Arg Asn Arg Gly Ser Gly
                350                 355                 360
Phe Gly His Asn Gly Val Asp Gly Asn Gly Val Gly Gln Ser Gln
```

```
                    365                     370                     375
Ala Gly Ser Gly Ser Thr Pro Ser Glu Pro His Pro Val Leu Glu
                    380                     385                     390
Lys Leu Arg Ser Ile Asn Asn Tyr Asn Pro Lys Asp Phe Asp Trp
                    395                     400                     405
Asn Leu Lys His Gly Arg Val Phe Ile Ile Lys Ser Tyr Ser Glu
                    410                     415                     420
Asp Asp Ile His Arg Ser Ile Lys Tyr Asn Ile Trp Cys Ser Thr
                    425                     430                     435
Glu His Gly Asn Lys Arg Leu Asp Ala Ala Tyr Arg Ser Met Asn
                    440                     445                     450
Gly Lys Gly Pro Val Tyr Leu Leu Phe Ser Val Asn Gly Ser Gly
                    455                     460                     465
His Phe Cys Gly Val Ala Glu Met Lys Ser Ala Val Asp Tyr Asn
                    470                     475                     480
Thr Cys Ala Gly Val Trp Ser Gln Asp Lys Trp Lys Gly Arg Phe
                    485                     490                     495
Asp Val Arg Trp Ile Phe Val Lys Asp Val Pro Asn Ser Gln Leu
                    500                     505                     510
Arg His Ile Arg Leu Glu Asn Asn Glu Asn Lys Pro Val Thr Asn
                    515                     520                     525
Ser Arg Asp Thr Gln Glu Val Pro Leu Glu Lys Ala Lys Gln Val
                    530                     535                     540
Leu Lys Ile Ile Ala Ser Tyr Lys His Thr Thr Ser Ile Phe Asp
                    545                     550                     555
Asp Phe Ser His Tyr Glu Lys Arg Gln Glu Glu Glu Glu Ser Val
                    560                     565                     570
Lys Lys Glu Arg Gln Gly Arg Gly Lys
                    575


<210> 49
<211> 139
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1447955CD1

<400> 49
Met Glu Thr Asp Cys Asn Pro Met Glu Leu Ser Ser Met Ser Gly
  1               5                  10                  15
Phe Glu Glu Gly Ser Glu Leu Asn Gly Phe Glu Gly Thr Asp Met
                 20                  25                  30
Lys Asp Met Arg Leu Glu Ala Glu Ala Val Val Asn Asp Val Leu
                 35                  40                  45
Phe Ala Val Asn Asn Met Phe Val Ser Lys Ser Leu Arg Cys Ala
                 50                  55                  60
Asp Asp Val Ala Tyr Ile Asn Val Glu Thr Lys Glu Arg Asn Arg
                 65                  70                  75
Tyr Cys Leu Glu Leu Thr Glu Ala Gly Leu Lys Val Val Gly Tyr
                 80                  85                  90
Ala Phe Asp Gln Val Asp Asp His Leu Gln Thr Pro Tyr His Glu
                 95                 100                 105
Thr Val Tyr Ser Leu Leu Asp Thr Leu Ser Pro Ala Tyr Arg Glu
                110                 115                 120
Ala Phe Gly Asn Ala Leu Leu Gln Arg Leu Glu Ala Leu Lys Arg
                125                 130                 135
Asp Gly Gln Ser


<210> 50
<211> 314
<212> PRT
<213> Homo sapiens
```

<220>
<221> misc_feature
<223> Incyte ID No: 1454689CD1

<400> 50
Met Ser Ala Ala Glu Ala Gly Gly Val Phe His Arg Ala Arg Gly
1 5 10 15
Arg Thr Leu Ala Ala Phe Pro Ala Glu Lys Glu Ser Glu Trp Lys
20 25 30
Gly Pro Phe Tyr Phe Ile Leu Gly Ala Asp Pro Gln Phe Gly Leu
35 40 45
Ile Lys Ala Trp Ser Thr Gly Asp Cys Asp Asn Gly Gly Asp Glu
50 55 60
Trp Glu Gln Glu Ile Arg Leu Thr Glu Gln Ala Val Gln Ala Ile
65 70 75
Asn Lys Leu Asn Pro Lys Pro Lys Phe Phe Val Leu Cys Gly Asp
80 85 90
Leu Ile His Ala Met Pro Gly Lys Pro Trp Arg Thr Glu Gln Thr
95 100 105
Glu Asp Leu Lys Arg Val Leu Arg Ala Val Asp Arg Ala Ile Pro
110 115 120
Leu Val Leu Val Ser Gly Asn His Asp Ile Gly Asn Thr Pro Thr
125 130 135
Ala Glu Thr Val Glu Glu Phe Cys Arg Thr Trp Gly Asp Asp Tyr
140 145 150
Phe Ser Phe Trp Val Gly Gly Val Leu Phe Leu Val Leu Asn Ser
155 160 165
Gln Phe Tyr Glu Asn Pro Ser Lys Cys Pro Ser Leu Lys Gln Ala
170 175 180
Gln Asp Gln Trp Leu Asp Glu Gln Leu Ser Ile Ala Arg Gln Arg
185 190 195
His Cys Gln His Ala Ile Val Leu Gln His Ile Pro Leu Phe Leu
200 205 210
Glu Ser Ile Asp Glu Asp Asp Asp Tyr Tyr Phe Asn Leu Ser Lys
215 220 225
Ser Thr Arg Lys Lys Leu Ala Asp Lys Phe Ile His Ala Gly Val
230 235 240
Arg Val Val Phe Ser Gly His Tyr His Arg Asn Ala Gly Gly Thr
245 250 255
Tyr Gln Asn Leu Asp Met Val Val Ser Ser Ala Ile Gly Cys Gln
260 265 270
Leu Gly Arg Asp Pro His Gly Leu Arg Val Val Val Val Thr Ala
275 280 285
Glu Lys Ile Val His Arg Tyr Tyr Ser Leu Asp Glu Leu Ser Glu
290 295 300
Lys Gly Ile Glu Asp Asp Leu Met Asp Leu Ile Lys Lys Lys
305 310

<210> 51
<211> 355
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1568009CD1

<400> 51
Met Pro Phe Leu Gly Gln Asp Trp Arg Ser Pro Gly Gln Asn Trp
1 5 10 15
Val Lys Thr Ala Asp Gly Trp Lys Arg Phe Leu Asp Glu Lys Ser
20 25 30
Gly Ser Phe Val Ser Asp Leu Ser Ser Tyr Cys Asn Lys Glu Val
35 40 45
Tyr Asn Lys Glu Asn Leu Phe Asn Ser Leu Asn Tyr Asp Val Ala
50 55 60

Ala Lys Lys Arg Lys Lys Asp Met Leu Asn Ser Lys Thr Lys Thr
65 70 75
Gln Tyr Phe His Gln Glu Lys Trp Ile Tyr Val His Lys Gly Ser
80 85 90
Thr Lys Glu Arg His Gly Tyr Cys Thr Leu Gly Glu Ala Phe Asn
95 100 105
Arg Leu Asp Phe Ser Thr Ala Ile Leu Asp Ser Arg Arg Phe Asn
110 115 120
Tyr Val Val Arg Leu Leu Glu Leu Ile Ala Lys Ser Gln Leu Thr
125 130 135
Ser Leu Ser Gly Ile Ala Gln Lys Asn Phe Met Asn Ile Leu Glu
140 145 150
Lys Val Val Leu Lys Val Leu Glu Asp Gln Gln Asn Ile Arg Leu
155 160 165
Ile Arg Glu Leu Leu Gln Thr Leu Tyr Thr Ser Leu Cys Thr Leu
170 175 180
Val Gln Arg Val Gly Lys Ser Val Leu Val Gly Asn Ile Asn Met
185 190 195
Trp Val Tyr Arg Met Glu Thr Ile Leu His Trp Gln Gln Gln Leu
200 205 210
Asn Asn Ile Gln Ile Thr Arg Pro Ala Phe Lys Gly Leu Thr Phe
215 220 225
Thr Asp Leu Pro Leu Cys Leu Gln Leu Asn Ile Met Gln Arg Leu
230 235 240
Ser Asp Gly Arg Asp Leu Val Ser Leu Gly Gln Ala Ala Pro Asp
245 250 255
Leu His Val Leu Ser Glu Asp Arg Leu Leu Trp Lys Lys Leu Cys
260 265 270
Gln Tyr His Phe Ser Glu Arg Gln Ile Arg Lys Arg Leu Ile Leu
275 280 285
Ser Asp Lys Gly Gln Leu Asp Trp Lys Lys Met Tyr Phe Lys Leu
290 295 300
Val Arg Cys Tyr Pro Arg Lys Glu Gln Tyr Gly Asp Thr Leu Gln
305 310 315
Leu Cys Lys His Cys His Ile Leu Ser Trp Lys Gly Thr Asp His
320 325 330
Pro Cys Thr Ala Asn Asn Pro Glu Ser Cys Ser Val Ser Leu Ser
335 340 345
Pro Gln Asp Phe Ile Asn Leu Phe Lys Phe
350 355

<210> 52
<211> 179
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1677811CD1

<400> 52
Met Gln Arg Gln Asn Phe Arg Pro Pro Thr Pro Pro Tyr Pro Gly
1 5 10 15
Pro Gly Gly Gly Gly Trp Gly Ser Gly Ser Ser Phe Arg Gly Thr
20 25 30
Pro Gly Gly Gly Gly Pro Arg Pro Pro Ser Pro Arg Asp Gly Tyr
35 40 45
Gly Ser Pro His His Thr Pro Pro Tyr Gly Pro Arg Ser Arg Pro
50 55 60
Tyr Gly Ser Ser His Ser Pro Arg His Gly Gly Ser Phe Pro Gly
65 70 75
Gly Arg Phe Gly Ser Pro Ser Pro Gly Gly Tyr Pro Gly Ser Tyr
80 85 90
Ser Arg Ser Pro Ala Gly Ser Gln Gln Gln Phe Gly Tyr Ser Pro
95 100 105
Gly Gln Gln Gln Thr His Pro Gln Gly Ser Pro Arg Thr Ser Thr

110 115 120
Pro Phe Gly Ser Gly Arg Val Arg Glu Lys Arg Met Ser Asn Glu
125 130 135
Leu Glu Asn Tyr Phe Lys Pro Ser Met Leu Glu Asp Pro Trp Ala
140 145 150
Gly Leu Glu Pro Val Ser Val Val Asp Ile Ser Gln Gln Tyr Ser
155 160 165
Asn Thr Gln Thr Phe Thr Gly Lys Lys Gly Arg Tyr Phe Cys
170 175

<210> 53
<211> 403
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1756951CD1

<400> 53
Met Ala Ala Phe Leu Lys Met Ser Val Ser Val Asn Phe Phe Arg
1 5 10 15
Pro Phe Thr Arg Phe Leu Val Pro Phe Thr Leu His Arg Lys Arg
20 25 30
Asn Asn Leu Thr Ile Leu Gln Arg Tyr Met Ser Ser Lys Ile Pro
35 40 45
Ala Val Thr Tyr Pro Lys Asn Glu Ser Thr Pro Pro Ser Glu Glu
50 55 60
Leu Glu Leu Asp Lys Trp Lys Thr Thr Met Lys Ser Ser Val Gln
65 70 75
Glu Glu Cys Val Ser Thr Ile Ser Ser Ser Lys Asp Glu Asp Pro
80 85 90
Leu Ala Ala Thr Arg Glu Phe Ile Glu Met Trp Arg Leu Leu Gly
95 100 105
Arg Glu Val Pro Glu His Ile Thr Glu Glu Glu Leu Lys Thr Leu
110 115 120
Met Glu Cys Val Ser Asn Thr Ala Lys Lys Lys Tyr Leu Lys Tyr
125 130 135
Leu Tyr Thr Lys Glu Lys Val Lys Lys Ala Arg Gln Ile Lys Lys
140 145 150
Glu Met Lys Ala Ala Ala Arg Glu Glu Ala Lys Asn Ile Lys Leu
155 160 165
Leu Glu Thr Thr Glu Glu Asp Lys Gln Lys Asn Phe Leu Phe Leu
170 175 180
Arg Leu Trp Asp Arg Asn Met Asp Ile Ala Met Gly Trp Lys Gly
185 190 195
Ala Gln Ala Met Gln Phe Gly Gln Pro Leu Val Phe Asp Met Ala
200 205 210
Tyr Glu Asn Tyr Met Lys Arg Lys Glu Leu Gln Asn Thr Val Ser
215 220 225
Gln Leu Leu Glu Ser Glu Gly Trp Asn Arg Arg Asn Val Asp Pro
230 235 240
Phe His Ile Tyr Phe Cys Asn Leu Lys Ile Asp Gly Ala Leu His
245 250 255
Arg Glu Leu Val Lys Arg Tyr Gln Glu Lys Trp Asp Lys Leu Leu
260 265 270
Leu Thr Ser Thr Glu Lys Ser His Val Asp Leu Phe Pro Lys Asp
275 280 285
Ser Ile Ile Tyr Leu Thr Ala Asp Ser Pro Asn Val Met Thr Thr
290 295 300
Phe Arg His Asp Lys Val Tyr Val Ile Gly Ser Phe Val Asp Lys
305 310 315
Ser Met Gln Pro Gly Thr Ser Leu Ala Lys Ala Lys Arg Leu Asn
320 325 330
Leu Ala Thr Glu Cys Leu Pro Leu Asp Lys Tyr Leu Gln Trp Glu
335 340 345

```
Ile Gly Asn Lys Asn Leu Thr Leu Asp Gln Met Ile Arg Ile Leu
                350                 355                 360
Leu Cys Leu Lys Asn Asn Gly Asn Trp Gln Glu Ala Leu Gln Phe
                365                 370                 375
Val Pro Lys Arg Lys His Thr Gly Phe Leu Glu Ile Ser Gln His
                380                 385                 390
Ser Gln Glu Phe Ile Asn Arg Leu Lys Lys Ala Lys Thr
                395                 400


<210> 54
<211> 163
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1833547CD1

<400> 54
Met Asp Val Arg Ser Arg Val Asp Ser Lys Thr Leu Thr Arg Asn
 1              5                   10                  15
Thr Arg Ile Ile Ala Glu Ala Leu Thr Arg Val Ile Tyr Asn Leu
                20                  25                  30
Thr Glu Lys Gly Thr Pro Pro Asp Met Pro Val Phe Thr Glu Gln
                35                  40                  45
Met Ile Gln Gln Glu Gln Leu Asp Ser Val Met Asp Trp Leu Thr
                50                  55                  60
Asn Gln Pro Arg Ala Ala Gln Leu Val Asp Lys Asp Ser Thr Phe
                65                  70                  75
Leu Ser Thr Leu Glu His His Leu Ser Arg Tyr Leu Lys Asp Val
                80                  85                  90
Lys Gln His His Val Lys Ala Asp Lys Arg Asp Pro Glu Phe Val
                95                  100                 105
Phe Tyr Asp Gln Leu Lys Gln Val Met Asn Ala Tyr Arg Val Lys
                110                 115                 120
Pro Ala Val Phe Asp Leu Leu Leu Ala Val Gly Ile Ala Ala Tyr
                125                 130                 135
Leu Gly Met Ala Tyr Val Ala Val Gln His Phe Ser Leu Leu Tyr
                140                 145                 150
Lys Thr Val Gln Arg Leu Leu Val Lys Ala Lys Thr Gln
                155                 160


<210> 55
<211> 110
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1833723CD1

<400> 55
Met Gln Ala Ala Leu Glu Val Thr Ala Arg Tyr Cys Gly Arg Glu
 1              5                   10                  15
Leu Glu Gln Tyr Gly Gln Cys Val Ala Ala Lys Pro Glu Ser Trp
                20                  25                  30
Gln Arg Asp Cys His Tyr Leu Lys Met Ser Ile Ala Gln Cys Thr
                35                  40                  45
Ser Ser His Pro Ile Ile Arg Gln Ile Arg Gln Ala Cys Ala Gln
                50                  55                  60
Pro Phe Glu Ala Phe Glu Glu Cys Leu Arg Gln Asn Glu Ala Ala
                65                  70                  75
Val Gly Asn Cys Ala Glu His Met Arg Arg Phe Leu Gln Cys Ala
                80                  85                  90
Glu Gln Val Gln Pro Pro Arg Ser Pro Ala Thr Val Glu Ala Gln
```

95 100 105
Pro Leu Pro Ala Ser
110

<210> 56
<211> 1929
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1841446CB1

<400> 56
cggaagcaaa ggagccaaga ccatggcgaa agccggggat aagagcagca gcagcgggaa 60
gaaaagtcta aaacggaaag ccgctgccga agaacttcag gaggctgcag gcgctgggga 120
tggggcgacg gaaaacgggg tccaaccccc gaaagcggct gcctttccgc caggctttag 180
catttcggag attaaaaaca aacagcggcg acacttaatg ttcacgcggt ggaaacagca 240
gcagcggaag gaaaagttgg cagctaagaa aaaacttaaa aaagaaagag aggctcttgg 300
cgataaggct ccaccaaagc ctgtacccaa gaccattgac aaccagcgag tgtatgatga 360
aaccacagta gaccctaatg atgaagaggt cgcttatgat gaagctacag atgaatttgc 420
ttcttacttc aacaaacaga cttctcccaa gattctcatc acaacatcag atagacctca 480
tgggagaaca gtacgactct gtgaacagct ctccacagtt ataccaaact cacatgttta 540
ttacagaaga ggactggctc tgaaaaaaat tattccacag tgcatcgcaa gagatttcac 600
agacctgatt gttattaatg aagatcgtaa aaccccaaat ggacttattt tgagtcactt 660
gccaaatggc ccaactgctc attttaaaat gagcagtgtt cgtcttcgta aagaaattaa 720
gagaagaggc aaggacccca cagaacacat acctgaaata attctgaata attttacaac 780
acggctgggt cattcaattg gacgtatgtt tgcatctctc tttcctcata atcctcaatt 840
tatcggaagg caggttgcca cattccacaa tcaacgggat tacatattct tcagatttca 900
cagatacata ttcaggagtg aaaagaaagt gggaattcag gaacttggac cacgttttac 960
cttaaaatta aggtctcttc agaaaggaac ctttgattct aaatatggag agtatgaatg 1020
ggtccataag ccccgggaaa tggatacaag tagaagaaaa ttccatttat aaagtactga 1080
gagaatgata ttggattttg ctgaacaggc ctatcttgaa ctttggtaaa ttatttttga 1140
cagaatactc ttttcaaaat ggcatttgct gatttcataa acctttcacg tctggacgaa 1200
ttaccaaatg ccatgaattg ccactgtgtg tttatgtaga aaatacaaat aaaagttatt 1260
ttgatggctt aggtttcctt aaacttagtt ctcttgtttt tgggtaactg tgaataatta 1320
agttggaatc aagattcaga ttaactttcc tatttgcata gaacacatga gaggaataaa 1380
atggttggta aatattggct aacccttgat ttttatacca gattaacctt ggattcccag 1440
tgtctggcac agttttaata gcttaaatgg aggccaggtt tctggatgtt ttaacattct 1500
cttaagcctt cagaagggta aaaaatttaa agcaaaatga tctaccaggg tttaaagcaa 1560
agttgcaaat tactgaagct aatctttgct tcctgatttt gaggttttttg gttttttgtg 1620
cccacgttgt ggggagctct tttttacctc attacatggt gctgtagtac tccattcagg 1680
cactgaaaca aagttaaccc tataagtaac tcatggatgg aaaccccgtag aacttaacag 1740
cctcctcctg accttaaaag aataaaggtt cacagtttac ctttaattcc ctagcagtct 1800
tgccagatgt atggcataaa gtcatgtgag aagagtaggt ggaaaaaact gtacaaactt 1860
aacccccttca ggtgttcaga acagattaat ataccatgta tttaatacca ataataatgc 1920
aaaataaag 1929

<210> 57
<211> 2113
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1850310CB1

<400> 57
ctgccacgcc actgtgtcat gaaggaaagt gaaagggaac gaggaagtag gaatgcccac 60
gctcgttgac tccgtgggtg aatacagcag ttaggacata cacaccatca cctttgaaag 120
tgcttgtttg ggggagggaa ggacatacgg gtaactagaa ctacccagcg agtcgtccag 180
aggagaggat caggtttgag tcaggaggct ccctgtactg gagtcgtccc actattcctc 240
aagaaatctt agaaccagct tgtgaggaaa aacatttttt aatgtaataa aaatatgcca 300
ttattctttg aaatgccaaa tgatataaat attttgccta atacatattt attgtagatg 360
aaatgcactc ttctcgatga ggcctcgatt tgaatcaatg gggtgggcca caggaaatgt 420
cagaggaacc agaactcaga actcttcctc ctggaccttt cttcccttcc cttggaggta 480

WEST

tcccttttgaa tcaggcctct ctcttctcat cagtctgtag cttcccccct tgtataacct 540
gctttccttt ttacatttat taaaagtgga ttttgtaaaa gcatttcatt gacacgcgac 600
ctatcacaga caatggaatt cgtcagtggt ggtaagactg aaatcctgat gcttttcaca 660
cttcttgtct cttgctatgt atttctgcct ctagccttgc catgttttgc cttttttttt 720
tctttttggc caattccttt ttatatgtgc ccacaacaga ggtggggaga cacggagcac 780
cctgggtcct tcccagcgct gctgggcagg ccccgtctcc aggccccagc tgttgaaact 840
ttgaagggca acaaacaacc atccacactg ccggacccta ggctgttcag ggaggcagct 900
catttccacc ccggccccag gacacccagc ctgtgcccca caaggatctc tctaaatggg 960
agggattgag gctacttttc tgccaagccc tattaagtag taatgtgggg aaacccactg 1020
tgtcagtgca ggaagcccta gacaaatgtt ttcaaataaa tttcactgcc cagcctgcac 1080
agatttccat ttgaagtact tcccatccac cctgacaccc aaaggggttt ttttgttttg 1140
ttttgttttt gagacagggt cttgctttgt tgcccaggct ggagtgcagt gacgtggtca 1200
tagctcactg cagcctcaac ctcctgggct caagtgaccc tcctgcctca gcctcccaaa 1260
gttctgagat gataggcatg agccattgtg cctagcctat tttgattttt ttcttagagt 1320
caaggtcttg ctctgttgcc caggctgatc ttggacttgc gagccaccat gcctggctgg 1380
gtttttttaa aaatagaatc tcactgatag cctgcaagaa acagatgcag tgcctgcttc 1440
cgtatcagtc caaggagccc tcgtgtttgc cacctttacc tttgaacctc cccctgcctc 1500
cctgcctgtg tccgcttttg cagctcaatg cagccatgac aaggaaagaa aagacaaagg 1560
aaggccagag agccgcgcag ttctctgcag gtgcagatgc aggcagtgga ggtggcctga 1620
gcaggcagaa ggacaccaag cgccctatgt tgcttgtcat tcatgacgtg gtcttggagc 1680
ttctgactag ttcagactgc cacgccaacc ccagaaaata ccccacatgc cagaaaagtg 1740
aagtcctagg tgtttccatc tatgtttcaa tctgtccatc taccaggcct cgcgataaaa 1800
acaaaacaaa aaaacgctgc caggttttag aagcagttct ggtctcaaaa ccatcaggat 1860
cctgccacca gggttctttt gaaatagtac cacatgtaaa agggaatttg gctttcactt 1920
catctaatca ctgaattgtc aggctttgat tgataattgt agaaataagt agccttctgt 1980
tgtgggaata agttataatc agtattcatc tcttttgtttt ttgtcactct tttctctcta 2040
attgtgtcat ttgtactgtt tgaaaaatat ttcttctata aaattaaact aacctgcctt 2100
aagaaaaaaa aaa 2113

<210> 58
<211> 1652
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1887020CB1

<400> 58
gacgaggtgc tattcgagtg ctagtggtgg aacgggacca cacgtattca caggcctcca 60
ctgggctctc agtaggtggg atttgtcagc agttctcatt gcctgagaac atccagctct 120
ccctcttttc agccagcttt ctacggaaca tcaatgagta cctggccgta gtcgatgctc 180
ctcccctgga cctccggttc aacccctcgg gctacctctt gctggcttca gaaaaggatg 240
ctgcagccat ggagagcaac gtgaaagtgc agaggcagga gggagccaaa gtttctctga 300
tgtctcctga tcagcttcgg aacaagtttc cctggataaa cacagaggga gtggctttgg 360
cgtcttatgg gatggaggac gaaggttggt ttgacccctg gtgtctgctc caggggcttc 420
ggcgaaaggt ccagtccttg ggagtccttt tctgccaggg agaggtgaca cgttttgtct 480
cttcatctca acgcatgttg accacagatg acaaagcggt ggtcttgaaa aggatccatg 540
aagtccatgt gaagatggac cgcagcctgg agtaccagcc tgtggaatgc gccattgtga 600
tcaacgcagc cggagcctgg tctgcgcaaa tcgcagcact ggctggtgtt ggagaggggc 660
cgcctggcac cctgcagggc accaagctac ctgtggagcc gaggaaaagg tatgtgtatg 720
tgtggcactg cccccaggga ccaggcctag agactccgct tgttgcagac accagtggag 780
cctatttcg ccgggaagga ttaggtagca actacctagg tggtcgtagc ccccactgagc 840
aggaagaacc ggacccggcg aacctggaag tggaccatga tttcttccag gacaaggtgt 900
ggccccattt ggccctgagg gtcccagctt ttgagactct gaaggttcag agcgcctggg 960
ccggctatta cgactacaac acctttgacc agaatggcgt ggtgggcccc cacccgctag 1020
ttgtcaacat gtactttgct actggcttca gtggtcacgg gctccagcag gcccctggca 1080
ttgggcgagc tgtagcagag atggtactga agggcaggtt ccagaccatc gacctgagcc 1140
ccttcctctt tacccgcttt tacttgggag agaagatcca ggagaacaac atcatctgag 1200
catgtgtgct ctgcactggc tccactggct tgcatcctgg ctgtgttcac agccttgttt 1260
gctgcttcca tcttccccag tactgtgcca ggccttctcc ccctccccag tgtcctctcc 1320
tctcaggcag gccattgcac ccatatggct gggcaggcac aggcagtgag gccgaggcca 1380
atagcgagtg atgagcggga tcctaggact gatctgtagc ccatgctgat gtcacccacc 1440
agggcaatcc atctggaggc ctgagcaccc tggcccagga ctggcttcat cctggcactg 1500
accaggaaag actgcctctg accctcttag cagacagagc ccaggcatgg gagcactctg 1560
gggcagcctg gctcaggttt attgattttc gtctgtttac cctatccatt aatcaataca 1620
tgtaattaac tccttccaaa aaaaaaaaaa aa 1652

WEST

<210> 59
<211> 1120
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1911421CB1

<400> 59
agagcgccgg aagcggtccg agaatgaaga gtgtgatcta ccatgcattg tctcagaaag 60
aggcgaatga ctccgatgtc cagccttcag gagcacagcg ggccgaggcc ttcgtgaggg 120
ccttcctgaa gcgcagcacg ccccgcatga gcccgcaggc ccgcgaggac cagctgcagc 180
gcaaggcggt ggtcctggag tacttcaccc gccacaagcg caaggagaag aagaagaaag 240
ccaaaggcct ctctgccagg caaaggaggg agctgcggct ctttgacatt aaaccagagc 300
agcagagata cagccttttc ctccctctcc atgaactctg gaaacagtac atcagggacc 360
tgtgcagtgg gctcaagcca gacacgcagc cacagatgat tcaggccaag ctcttaaagg 420
cagatcttca cggggctatt atttcagtga caaaatccaa atgcccctct tatgtgggta 480
ttacaggaat ccttctacag gaaacaaagc acattttcaa aattatcacc aaagaagacc 540
gcctgaaagt tatcccccaag ctaaactgcg tgttcactgt ggaaaccgat ggctttattt 600
cctacattta cgggagcaaa ttccagcttc ggtcaagtga acggtctgcg aagaagttca 660
aagcgaaggg aacgattgac ctgtgaattc tttgccgtct aaggcagttg tttatgacag 720
ctgaaaactg gacactccct aaatgtccac ctttcagtga agagatagtt aagccaattc 780
catttataga ccacctccag ccagtgacgc tccgagttga ggatgttgaa caacatggga 840
aggtcgcagc gtactaagtg aagaagtcag aggacagagg aatttctctt tctaggagat 900
tttcattttg tgtgactccc atggggagga acagactggc aggaagcaca ccggggttaa 960
cactggttga cttgaatagg attattcgat ttttaaaaat acttttccat gttttctgag 1020
tgctctatga taaatcagtt gcatctgtga taatacagta catatgtgga cataaacagg 1080
gatcaaataa aggaggtatt gctgcaaaaa aaaaaaaaaa 1120

<210> 60
<211> 1000
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1911910CB1

<400> 60
agtcctggaa agcgttgttg gcccggttgc tctggagccg ggtctcgggt ctggtggctg 60
ccggccctgc ggcatctcgc catggggagc acggagagca gcgagggccg cagggtgtcc 120
ttcggagtgg acgaggagga gcgggtccgg gtgctgcagg gtgtccggct gtctgaaaac 180
gtggtgaacc gcatgaagga gcccagctct ccaccccctg ctcccacatc ttctaccttt 240
ggccttcaag atggcaactt gagagcccct cacaaagaat ccacactgcc caggtcgggg 300
agcagtggtg gccagcagcc ctcagggatg aaggagggtg tcaagaggta tgaacaggag 360
catgctgcta tccaggataa gctcttccag gtggcaaaga gggaaagaga ggctgccacc 420
aagcactcca aggcatccct gcccacgggc gaaggcagca tcagccatga ggagcagaag 480
tcagtccggc tggccaggga gctggagagc agagaggcag agctaagacg ccgtgacacc 540
ttctacaagg agcagctgga gcgtattgag aggaagaatg ctgagatgta taaactgtct 600
tcagagcaat tccatgaggc agcctcaaag atggagagca caataaagcc ccgcagggtg 660
gagcccgtct gctcagggtt gcaggcccag attctccact gctaccgaga tcgcccgcat 720
gaggtgctgc tgtgctcgga cctggtcaag gcataccagc gctgcgtgag cgccgcccac 780
aagggctgag gagcagacat cattccctgc cctggcagtg acttggagcc ctgaagaagg 840
gaccaatcat gggaccacag ccactgtgcc ctgccgtttc ctgctgggcc cctgcatatg 900
cccctgagcc tggggctgcc acgtgtttag gaaacaaagt atgcgctact gtctgaaaac 960
aaataaagca gatgcctttg ttttcaaaaa aaaaaaaaaa 1000

<210> 61
<211> 2273
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1928920CB1

<400> 61
gggggtgaag cgatacgttt tgcccgcatt cggggcgcgc ggactggggg ggtccctgtg 60
gggctcccgg agttaagatg gcgtcctcag cggaggggga cgaggggact gtggtggcgc 120
tggcgggggt tctgcagtcg ggtttccagg agctgagcct taacaagttg gcgacgtccc 180
tgggcgcgtc agaacaggcg ctgcggctga tcatctccat cttcctgggt taccccttttg 240
ctttgtttta tcggcattac cttttctaca aggagaccta cctcatccac ctcttccata 300
cctttacagg cctctcaatt gcttatttta actttggaaa ccagctctac cactccctgc 360
tgtgtattgt gcttcagttc ctcatccttc gactaatggg ccgcaccatc actgccgtcc 420
tcactacctt ttgcttccag atggcctacc ttctggctgg atactattac actgccaccg 480
gcaactacga tatcaagtgg acaatgccac attgtgttct gactttgaag ctgattggtt 540
tggctgttga ctactttgac ggagggaaag atcagaattc cttgtcctct gagcaacaga 600
aatatgccat acgtggtgtt ccttccctgc tggaagttgc tggtttctcc tacttctatg 660
gggccttctt ggtagggccc cagttctcaa tgaatcacta catgaagctg gtgcagggag 720
agctgattga cataccagga aagataccaa acagcatcat tcctgctctc aagcgcctga 780
gtctgggcct tttctaccta gtgggctaca cactgctcag cccccacatc acagaagact 840
atctcctcac tgaagactat gacaaccacc ccttctggtt ccgctgcatg tacatgctga 900
tctggggcaa gtttgtgctg tacaaatatg tcacctgttg gctggtcaca gaaggagtat 960
gcattttgac gggcctgggc ttcaatggct ttgaagaaaa gggcaaggca aagtgggatg 1020
cctgtgccaa catgaaggtg tggctctttg aaacaaaccc ccgcttcact ggcaccattg 1080
cctcattcaa catcaacacc aacgcctggg tggcccgcta catcttcaaa cgactcaagt 1140
tccttggaaa taaagaactc tctcagggtc tctcgttgct attcctggcc ctctggcacg 1200
gcctgcactc aggatacctg gtctgcttcc agatggaatt cctcattgtt attgtggaaa 1260
gacaggctgc caggctcatt caagagagcc ccaccctgag caagctggcc gccattactg 1320
tcctccagcc cttctactat ttggtgcaac agaccatcca ctggctcttc atgggttact 1380
ccatgactgc cttctgcctc ttcacgtggg acaaatggct taaggtgtat aaatccatct 1440
atttccttgg ccacatcttc ttcctgagcc tactattcat attgccttat attcacaaag 1500
caatggtgcc aaggaaagag aagttaaaga agatggaata atccatttcc ctggtggcct 1560
gtgcgggact ggtgcagaaa ctactcgtct cccttttcac agcactcctt tgccccagag 1620
cagagaatgg aaaagccagg gaggtggaag atcgatgctt ccagctgtgc ctctgctgcc 1680
agccaagtct tcatttgggg ccaaagggga aacttttttt tggagaaggc gtcttgcttt 1740
gtcacccacg ctggaatgca gtggcgggat ctcagctcac cgcaacctcc acctcctggg 1800
ttcaagtgat ttcctgcct cagcctccca agtagctggg aatacaggca cgccaccatg 1860
cccagctaat ttttgtattt tcagtagaaa cgggatttca ccacgttggc caggctggtc 1920
tcgaactcct gaccgcaagt gatccacccg cctccgcctc ccaaagtgct gggattacag 1980
gcgtgagcca ccgtgcccgg cccaaagggg aaactcttgt gggaggagca gaggggctca 2040
catctcccct ctgattcccc catgcacatt gccttatctc tccccatcta gccaggaatc 2100
tattgtgttt ttcttctgcc aatttactat gattgtgtat gtgccgctac caccaccccc 2160
cccatggggg ggtggagagg ggtgcaaggc cctgcctgct ccactttttc taccttggaa 2220
ctgtattaga taaaatcact tctgtttgtt cagtttttca aaaaaaaaaa aaa 2273

<210> 62
<211> 925
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2170846CB1

<400> 62
cgcagaaaca gcctatagac gccacgagtc ggcggcgcta ccgaggggct gtgggcgcgc 60
agctggaacc tccggctgtc agtgcgctta cagttcctaa ccccgaccct gcgcgcaccc 120
gcactatggc agccccgccg cagctaaggg ctctgctcgt agtcgtcaac gcactgctgc 180
gcaagcgccg ctaccacgct gcgttggccg tgcttaaggg cttccggaac ggggctgtct 240
atggagccaa aatccgggcc cctcacgcgc tggtcatgac ctttctcttc cggaatggca 300
gcctccagga gaagctgtgg gccatactgc aggccacata tatccactcc tggaacctgg 360
cacggtttgt gttcacctac aagggtctcc gtgccctgca gtcctacata caaggcaaga 420
cctacccagc acacgcattc ctggcggcct tcctcggggg tatcctggtg tttggagaaa 480
acaataacat caacagccag atcaacatgt acctgttgtc acgcgtcctg tttgccctga 540
gccgcctggc tgtagagaag ggctacatcc ctgaacccag gtgggacccg ttcccgctgc 600
tcactgcggt ggtgtggggg ctggtgctgt ggctctttga gtatcaccga tccaccctgc 660
agccctcgct gcagtcctcc atgacctacc tctatgagga cagcaatgta tggcacgaca 720
tctcagactt cctcatctat aacaagagcc gtccctccaa ttaatgcagc cctgaggtgt 780
ctggctgtgg ctcaagattt ggccccatgc agaccctccc aaaggatact gccttctcaa 840
gatcataggc ctcagactcc aactggtgtt atcccagggt tctgtttgct gaagtaaaaa 900
cactgatttt aaaaaaaaaa aaaaa 925

WEST

<210> 63
<211> 1570
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2176361CB1

<400> 63
```
gggcctccgt ctccgacgct gacttcctca gcgcaccagc actgtctccg gtggcaagtc   60
gaccagctcc tcccacacgg attctgtgta gaaatccacg gtatgtgcat tggaaatgga  120
cagggcatcc ctcaggaact gcagcagtcc ctgcaacttg gcacgcagcg tgggcaggtc  180
cggggtcacc gggagagggc aagaagccgc catgacgctc accctccgag gctgtaggtg  240
gcgcaaacat ggccgtgagg cgccagctcg gcccaaatta gcacaggcag gctcagtaga  300
gctccgagcc ggattccttc tgagcgattg attgcttcgg cttggtgacg tattttgtgg  360
gcgtccgtcg tggtcttctg atgacgcatt agtcggctgc aatggcgccg gtgaggcggt  420
ccgcgaagtg gcggcctggt ggtattgagg cgcgtggtga aggggtttcc actgtcgggt  480
acaggaataa gaatgtgaga cagaagacat ggcggcctaa ccacccgcaa gccttcgtgg  540
ggagcgttcg cgagggacaa ggctttgctt ttcgaagaaa actgaaaata cagcaaagtt  600
acaagaaatt gctacggaag gaaaagaagg ctcaaacgtc actggaatct caattcacag  660
atcgataccc agataatctg aaacatctct atttagctga agaggaaaga cataggaagc  720
aagcaagaaa agtcgaccat cctttgtcag aacaagttca ccagccgttg cttgaagaac  780
agtgtagcat tgacgagcct ttatttgaag atcagtgtag ctttgaccag cctcagccag  840
aagaacaatg tattaaaaca gtaaactcct ttacaattcc aaagaaaaat aaaaagaaaa  900
catcaaatca aaaagcacaa gaagaatatg aacagataca agccaaacgt gctgctaaga  960
aacaagaatt cgagaggaga aaacaggaga gagaagaagc ccaaaggcag tacaaaaaga 1020
agaaaatgga agtgtttaaa atactgaaca aaaagactaa aaagggccaa ccaaacttga 1080
atgtacaaat ggagtacctt cttcaaaaaa tacaagaaaa atgttaaaca ttttgttcct 1140
acaggttaaa atatctgctg cctattaggt tcttctgtga catgtgcctc ccagcagtga 1200
actaaatttg tcgacataaa ctggattgct aaactatgct aaatataaga tgttcacata 1260
tttttattat ggtaaaaaat tttctaaata tgttctacat gtttcttatt tatttgcctc 1320
tgaaggaagg ttggcctgaa gaactgaaag aacctcttat tttgcaagac aggcccaagc 1380
atgtaatact tttgtaccat atgagattta tatgaaataa attttttaaa aataaggaat 1440
cagagctatc aatgaagcat ttcaatgaaa tatttcaatt tagtaaacag ttacgttgtt 1500
ttaaaattta ttttaatgat gagggaggca aagactcctc ttggactttt tatttatttt 1560
aagtacatgt                                                        1570
```

<210> 64
<211> 1868
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2212732CB1

<400> 64
```
tcaccatgag gatccagccc ttggtaaggc cgggactgca gaagcagtca tccctgaaaa   60
ccatgaggtt ctggcaggcc cagatgagca ccctcaggac acagatgcaa gagatgctga  120
tggggaggct agagatcggg agccagcaga ccaagctttg ctgcttagcc agtgtgggga  180
taaccttgag tcccctctgc ctgaagctag ctcagctcca ccggggccaa cccttgggac  240
actgcctgaa gtagagacaa taagggcatg ctccatgccc caggagcttc ctcagtcccc  300
caggacccga cagcctgagc cagatttcta ctgtgtcaag tggatccctt ggaaaggaga  360
acagacaccc atcatcaccc agagcactaa cggccccttgc cctctccttg ccatcatgaa  420
catcctcttt cttcagtgga aggtgaagct ccccccgcag aaggaagtga tcacatcgga  480
tgagctcatg gcccatcttg gaaactgcct cctgtccatc aagccccagg agaagtcaga  540
gggacttcag cttaattttc agcagaatgt ggatgatgca atgacagtgc tgcctaaact  600
ggccacaggt ctggatgtca atgtgcgatt cacaggcgtc tctgattttg agtatacacc  660
cgagtgcagt gtctttgacc tgctaggcat acctctgtac catggctggc ttgttgatcc  720
acagcagagt cctgaggctg tgcgtgcagt tgggaaactg agttacaacc agctggtgga  780
gaggatcatc acctgcaaac actccagtga caccaacctc gtgacagaag gcctgattgc  840
agagcagttc ctggagacca ccgcggccca gctgacctac cacggactgt gtgagctgac  900
agcagctgct aaggagggtg aacttagcgt ctttttccga aacaaccact ttagcaccat  960
gactaagcat aagagtcact tatacctact ggtcactgac cagggctttc tacaggagga 1020
gcaagtcgta tgggagagcc tgcacaatgt ggatggagac agctgctttt gtgactctga 1080
ctttcacctg agtcattccc tggcaaggg gcctggagca gaaggtggga gtggctcccc 1140
```

agaaaagcag ctgcaggtag accaggacta cctgattgct ctgtccctgc agcagcaaca 1200
gccacgaggc ccgctggggc ttaccgactt ggagctggcc cagcagcttc agcaagagga 1260
gtatcaacag cagcaggcag cgcagccagt gcggatgcgg acgcgggtcc tgtcactgca 1320
ggggagagga gccacatctg gacgcccagc cggggagcgt cggcagaggc cgaagcacga 1380
gtcagactgc attctgctgt agctctgccc cagtgccagg ctggcctgcc ccttcttcca 1440
gaggctatgg ctagttggct tgctccccg cctccacccc tgagatgtgc tggataactt 1500
atttatggac tgttggggat gagagcaggc aacaaatgcc aaggtcagac ttggtaatgt 1560
ccttgacctc acgtgctgct gccttctctg cctcccaccc agggcaacac taggattggt 1620
gggtttctgg ttctcaactc ccggtccctg aatagtcaca cgtatgtaca gactgaggct 1680
ctggggtgag gtccctatcc agaatgcatc tcttctgctt cccatccctg ctgcctggat 1740
gctcctgatc acctaggcag gcctgtctcc agttgtttca gagcttaatt tgggtttcta 1800
tctcttattt gtaatgcctt cctggggttt ggaaataaaa cttctggccg ggcaaaaaaa 1860
aaaaaaaa 1868

<210> 65
<211> 1401
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2303457CB1

<400> 65
ccaagctgca gctggcaggg attgcggggt gccggccgtc tgagtttttt taaaactgct 60
cgccgcgaag tctgtctgca gccaaaatgt ccaacagaaa caacaacaag cttcccagca 120
acctgccgca gttacagaat ctaatcaagc gagacccgcc ggcctacatc gaggagtttc 180
tacagcagta taatcactac aaatccaatg tggagatttt caaattgcaa ccaaataaac 240
ccagcaaaga actagcagag ctggtgatgt ttatggcaca gattagtcac tgctacccag 300
agtacctaag taattttcct caagaggtga aagatcttct ctcctgcaat cataccgtat 360
tggatccaga tctgcgaatg acattttgca aagctttgat cttgctgaga aataagaatc 420
tcatcaatcc atcaagcctg ctagaactct tctttgaact ttttcgttgc catgataaac 480
ttctgcgaaa gactttatac acacatattg tgactgatat caagaatata aatgcaaaac 540
acaagaacaa taaagtgaat gtagtattgc aaaatttcat gtacaccatg ttaagagata 600
gcaatgcaac cgcagccaag atgtctttag atgtaatgat tgaactctac agaaggaaca 660
tctggaatga tgcaaaaact gtcaatgtta tcacaactgc atgtttctct aaggtcacca 720
agatattagt tgccgctttg acattctttc ttgggaaaga tgaagatgaa aaacaggaca 780
gtgactccga atctgaggat gatggaccaa cagcaagaga cctgctagta caatatgcta 840
cagggaagaa aagttccaaa aacaagaaaa agttggaaaa ggcaatgaaa gtgctcaaga 900
aacaaaaaaa gaagaaaaaa ccagaggtgt ttaacttttc agccattcac ttgattcatg 960
atccccaaga ttttgcggaa aaactactaa agcagcttga gtgctgtaag gagaggtttg 1020
aagtgaagat gatgctcatg aaccttatct ccagattggt gggaattcat gagcttttcc 1080
tcttcaattt ctatcccttt ttgaaaaggt ttctgaagcc ccaccaaagg gaggtaacca 1140
agatccttct gtttgttgaa aaagattctc atcacttagt accccaaggg ttttttaatt 1200
catggttaat gcttggggaa aaaatttttt ttaacggaaa aaaatctggg aaaatgttaa 1260
tgacagttgg gaatttaatg gttaaaagag gggtatataa acgctccaaa gtgttccttg 1320
ggggaaatag tgttggaaga aatttttttc aaaaaaatcc agggggctca agttaaaaaa 1380
acccccccagg cgggtgtgta t 1401

<210> 66
<211> 1409
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2317552CB1

<400> 66
gtcggggctg cgccgtacaa cttccggctg taaagatggc ggcttcctag tgagtcggcg 60
gctgatttag aaggaggttc aggctacggt gagccgaagc cacacaggag ccatggaagt 120
ggcagagccc agcagcccca ctgaagagga ggaggaggaa gaggagcact cggcagagcc 180
tcggccccgc actcgctcca atcctgaagg ggctgaggac cgggcagtag ggcacaggc 240
cagcgtgggc agccgcagcg agggtgaggg tgaggccgcc agtgctgatg atgggagcct 300
caacacttca ggagccggcc ctaagtcctg gcaggtgccc ccgccagccc ctgaggtcca 360
aattcggaca ccaagggtca actgtccaga gaaagtgatt atctgcctgg acctgtcaga 420
ggaaatgtca ctgccaaagc tggagtcgtt caacggctcc aaaaccaacg ccctcaatgt 480

ctcccagaag atgattgaga tgttcgtgcg gacaaaacac aagatcgaca aaagccacga 540
gtttgcactg gtggtggtga acgatgacac ggcctggctg tctggcctga cctccgaccc 600
ccgcgagctc tgtagctgcc tctatgatct ggagacggcc tcctgttcca ccttcaatct 660
ggaaggactt ttcagcctca tccagcagaa aactgagctt ccggtcacag agaacgtgca 720
gacgattccc ccgccatatg tggtccgcac catccttgtc tacagccgtc caccttgcca 780
gccccagttc tccttgacgg agcccatgaa gaaaatgttc cagtgcccat atttcttctt 840
tgacgttgtt tacatccaca atggcactga ggagaaggag gaggagatga gttggaagga 900
tatgtttgcc ttcatgggca gcctggatac caagggtacc agctacaaat atgaggtggc 960
actggctggg ccagccctgg agttgcacaa ctgcatggcg aaactgttgg cccacccccct 1020
gcagcggcct tgccagagcc atgcttccta cagcctgctg gaggaggagg atgaagccat 1080
tgaggttgag gccactgtct gaaccatccc tgtacatctg caccttcttg tgcaaggaag 1140
tccttggcct aaagccttgg ttctcaaact gggttccttg ggacctccgg ggtgggggggg 1200
ttccaggagg cacgtagggt accttgcagg gtcctaggag ggaaacccag gattccagga 1260
gggatcccag gaactgtggg cacccatttt ctgtgtctcc cagcccattt ccactcctag 1320
tttgtcatgg ataatttttg ttcttccctg tgtgattttt gccatcaaaa taaaaatttg 1380
agactcgtta accgaaaaaa aaaaaaaaa 1409

<210> 67
<211> 1888
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2416366CB1

<400> 67
tgtgattcgg agagtgagga tccctcaagg aaccaggcta gtgattcggg aaatgaggag 60
ctactcaaac cccgagtcag tgactctgag agtgaggggc ctcagaaggg gcctgccagt 120
gactcagaaa ctgaggatgc gtccagacac aaacagaagc cagagtcaga tgatgacagc 180
gacagggaga ataagggaga ggatacagaa atgcagaatg actccttcca ttcagacagc 240
catatggaca gaaaaaagtt tcacagttct gatagtgagg aggaagaaca caaaaagcaa 300
aaaatggaca gtgatgaaga tgaaaaagag ggtgaggagg agaaagtagc gaagagaaaa 360
gctgctgtgc tttctgatag tgaagatgaa gagaaagcat cagcaaagaa gagtcgtgtt 420
gtctctgatg cagatgactc tgacagtgat gctgtatcag acaagtcagg caaaagagag 480
aagaccatag catctgacag tgaggaagaa gctgggaaag aattgtctga taagaaaaat 540
gaagagaagg atctgtttgg gagtgacagt gagtcaggca atgaagaaga aaatcttatt 600
gcagacatat ttggagaatc tggtgatgaa gaggaagaag aatttacagg ttttaaccaa 660
gaagatctgg aagaagaaaa aggtgaaaca caggtaaaag aagcagaaga ttcagattct 720
gatgataaca taaagagagg aaaacatatg gactttctgt cagattttga gatgatgttg 780
cagcgaaaaa agagcatgag tggcaagcgc agacggaacc gcgatggtgg caccttttatt 840
agtgatgcag acgacgtcgt gagtgccatg atcgtcaaga tgaatgaagc tgctgaggaa 900
gacagacagt tgaacaatca aaaaaagcca gcactgaaaa aattaacttt actgcctgct 960
gtagttatgc accttaagaa gcaggacctt aaagaaacat tcattgacag tggtgtgatg 1020
tctgccatca aagaatggct ctcacctcta ccagatagga gtttgcctgc actcaagatc 1080
cgggaggagc tgctgaagat cctgcaagag ctgcctagtg tgagccagga gaccctgaag 1140
catagtggga ttggacgagc agtgatgtat ctctataaac accccaagga gtcaaggtct 1200
aacaaggaca tggcagggaa attaatcaat gagtggtcta ggcctatatt tggtcttacc 1260
tcaaactaca aggaatgac aagagaagaa agggagcaga gagatctaga acagatgcct 1320
caacgacgaa gaatgaacag cactggtggt cagacaccca gaagagacct ggaaaaggtg 1380
ctgacaggag aggagaaggc tcttagacct ggagatcctg gattctgtgc ccgtgcaagg 1440
gtcccaatgc cttcaaacaa ggactatgtt gtcaggccca aatggaatgt ggaaatggag 1500
tcatccaggt ttcaggcgac ctccaagaag ggtatcagtc gactggataa acagatgaga 1560
aagttcacag atataaggaa aaaaagcaga tctgcacacg cagtgaaaat cagcattgag 1620
ggcaacaaaa tgccattgtg accttgcctg gaatgtgtcc ccatctctac tctaagaaat 1680
gcgcaatgga ctctttggag aaagaagata ttttaaaaca tttttagtgt gtctgtaaat 1740
ggttcagcgt gtatcagatg ttgtcatagg actcacattt ctctcagtta tatttaaaac 1800
cgttgtgtac tttgtacaaa ggaatactag tcatacttct ataaacttta cacaataaaa 1860
tttcattctg gttaaaaaaa aaaaaaaa 1888

<210> 68
<211> 1897
<212> DNA
<213> Homo sapiens

<220>
<221> unsure

WEST

<222> 1892
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte ID No: 2472980CB1

<400> 68
ttattgaaga catggccgcc gcgtattttc cagactgcat agtcagaccc tttggctcct 60
cagtcaacac ttttgggaag ttaggatgtg atttggacat gtttttggat ctagatgaaa 120
ccagaaacct cagcgctcac aagatctcag gaaattttct gatggaattt caagtgaaaa 180
atgttccttc agaaagaatt gcaactcaga agatcctgtc tgtgttagga gagtgccttg 240
accactttgg ccctggctgt gtgggtgtgc aaaaaatatt aaatgcccgg tgtccgctcg 300
tgaggttctc acaccaggcc tccggatttc agtgtgattt gactacgaac aataggattg 360
ccttgacaag ttccgaactc ctttatatat atggtgccct agactcaaga gtgagagcct 420
tggtgttcag tgtacggtgc tgggctcgag cacattcact aacaagtagt attcctggtg 480
catggattac aaatttctcc cttacaatga tggtcatctt ttttctccag agaagatcac 540
cccctattct tccaacacta gattccttaa aaaccctagc agatgcagaa gataaatgtg 600
taatagaagg caacaactgc acatttgttc gtgacttgag tagaattaaa ccttcacaga 660
acacagaaac attagaatta ctactgaagg aattttttga gtattttggc aattttgctt 720
tcgataaaaa ttccataaat attcgacagg gaagggagca aaacaaacct gattcttctc 780
ctctgtacat tcagaatcca tttgaaactt ctctcaacat aagcaaaaat gtaagtcaaa 840
gccagctgca aaaatttgta gatttggccc gagaaagtgc ctggatttta caacaggaag 900
atacagatcg accttccata tcaagtaatc ggccctgggg gctggtatcc ctattgctac 960
catctgctcc aaacagaaag tcctttacca agaagaaaag caataagttt gcaattgaaa 1020
cagtcaaaaa cttgctagaa tctttaaaag gtaacagaac agaaaatttc acaaaaacca 1080
gtgggaagag aacaattagt actcagacat gatggctgct acattgtgta aagaactggg 1140
cttagcctat caaatggtct gtggacttac ttggaaaaac tgatttgaaa ctttcacaga 1200
tctcagcttt catctgatgt cacttttcat gatcttctca ttggcccccct taacctggtc 1260
tgaagttctg ggatgttttc agtttgatca gtctgatact cagtggcact ttattaaaac 1320
atcagctgtg gagtgtggcg gtgcacacct gtagtcccag ctgctcagga ggctgaggca 1380
ggaggatctc ttgagcccag gattttgaat ccatcgtgga caacatagca agattccatc 1440
tctaaaaaaa atgaaaataa acataagcca caaggaatgg gtgaaagatt attgtaatgt 1500
gctttaacta aataggtaaa tatactaaac aaatgctaaa actcagtttt aggatgaaac 1560
cattgttgat atccacatca gtccctgttt agaaaacatt taaaatgact tttagttatg 1620
tacagtacgt tggcaatgaa tacattaagc ttcaaaattt ggtagtgctc tcgaatatgt 1680
atatttgtat ttttcaagcg aagttctctt attcacatat aaattaaagt gggttggtac 1740
tgatatcaaa aaatgtttat gtttttagaa cagacatttc agtcactgca ttcttaggta 1800
ttccaaacca aatatgatga catcaataga ttgcatttta aaaatattgt ttgatttttc 1860
tatggtcaaa aataaaattg tgttctactt tnctggt 1897

<210> 69
<211> 1132
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2541640CB1

<400> 69
gtggagcctc gaacccgaga cgctagaccc aatttggtgc ttatgggagg agtgggagtg 60
gcagaggcag cgagaccact cctttcttgg cccaccattt ctctcaccat ttttactgca 120
gtgaactcct ctcagggtgg ggggctggta cagaggcaac tccgctttca taactctcac 180
agggtccttt gtcggaggtg tccctgccct cctaccccag cgtggtggga gtgcgacgcc 240
aggcttctcc ctcctccctg gcctcctgtg cctccagcct ccacatctcc tgagatcctc 300
cccaccccac acttacacag gtccccacac gcccctggtg ccccctaaacc gccacccaac 360
cccactcacc ctggggccgg gactggggtg tcagagctga gccagggacc ttgggaggtg 420
gctgggaccg gcgcctcctg ctccctcttt cattttccct ttcggatttg gccaggatgg 480
aggacaggac aggatggctg agaggggcta gcgctccccc ttctcttaaa ggaacagggt 540
gaccctgctg gctggctccg cccttgggaa caggcacgga ccacgcccct tgctgggctc 600
tgtggccctc actctttagc taaatttcca tctggaaatg gggggagggg gtaaccagag 660
ccacatcctt tatcgaggaa gctgggggat tcctagaagc ccttcccaga cagaccctca 720
gtctggagcc tagatccact acgcactaac catgccctgt gaagatggtc ttcggggaca 780
gagggaggtt tggtgctgga gtttaagggg cagcaaacat ctaacagctt ctggttccag 840
ctctgccacc gctcaccttg gacaaatcat tcggcggtac tgagccaggc tttcctgcct 900
tgtcaacaaa ggcggccagt ccagctgacc cccaagtttc cttccagctt tacatctttt 960

WEST

ttggaacggg gtctctctac atagcccagg tcgatcttga actcctggct tcaagtgatc 1020
ctcacacctt agcctcctga gtagctggga ttctctactt cttaatattg aattcccaac 1080
catattacag agaaccaaaa aaacccttgt gcctgcaaat aaataaaaaa aa 1132

<210> 70
<211> 1763
<212> DNA
<213> Homo sapiens

<220>
<221> unsure
<222> 1612, 1613, 1737, 1745
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte ID No: 2695204CB1

<400> 70
aggctacagc tagggctgct cagtaactgc cccaaccact accacgtccc caagtgcaac 60
gggacccacg gaactacagg gtcctgctga gctcaggagc tgcacggaat ggtggcagtc 120
accttgccag tgcagcaggc atggatcaga ctgcaagcta ggggcacgag gcatcagttg 180
ggaagagggg acgcacagct aggcttgagg cctcttcatc gggatgtccc caggcccccc 240
agcccaggcc cagcataaag gccgtgttgg ggggcccccc tgacccaagg ggggcttcat 300
gcgccacgtg caggcggagc gtagtccatc ctcagagccg gacgtggccc ttcacagcct 360
ccagtcaggc agggggccct ccagggtggc ctgctcatgg gctacagccc agcagggggg 420
gcgacatccc ccggggtcta ccaggcccct gccccaccaa gcgaaggctg cttcctgctg 480
gagaagcccc agatgtcagc tctgaggaag aggggccagc ccctcggagg cgccggggat 540
ccctgggcca ccctactgct gccaacagtt ctgatgccaa agccacaccc ttctggagcc 600
acctgctgcc tgggcccaaa gagcctgttt tggacccaac agactgcggt cccatggggc 660
ggaggctgaa aggagcccgt cgcctgaagc tgagccccct tcgaagcctc cggaaggggc 720
caggcctgct gagccccccc agtgcctccc ctgttcctac ccctgctgtc agccgtaccc 780
tgctgggcaa ctttgaggaa tcattgctgc gaggacgttt tgcaccatct ggccacattg 840
agggcttcac agcagaaatt ggagctagtg ggtcatactg cccccagcac gtcacgctgc 900
ctgtcactgt cacattcttt gatgtttctg agcaaaatgc cccggctccc ttcctgggca 960
tcgtggatct gaacccccttg gggaggaagg gttacagcgt gcccaaggtg ggcaccgtcc 1020
aagtgacctt atttaacccc aaccagactg tggtaaagat gttccttgtg acctttgact 1080
tctcggacat gcctgctgcc cacatgacct tcctgcgcca tcgcctcttt ttggtgcctg 1140
tgggtgagga gggaaatgct aaccccaccc accgcctcct ctgctacttg ctgcacctca 1200
ggttccggag ctcccgctca ggccgcttaa gcctgcatgg agatatccgc ctgctttttt 1260
cccgccggag cctggagctg gacacagggc tcccctacga actgcaggct gtgaccgagg 1320
cccctcataa tccacgttat tcacctttgc cctgattgcc agcactctga acccatgcgg 1380
gctaatgacc tgcccatcct gctccatctt agagaacata tatggagaga cagcaagaga 1440
cccttcaggc ttgaattaaa gccctcacca tgctcacgcc caaatggatt atttgggtgt 1500
ttaaagcttc tgattcttac tacaccctgc cctacttcgg gtactccatg tgcctgtccc 1560
ctcccttggg tttcccagga cagcttaggt agtagggagg aactggagct annctgagat 1620
tttctcaagt tcccaggcaa gttatgccag cgttgcctcc ctgtcctggg caggggccac 1680
ttgttatttt atttattttt aatttataat ttattcaatt ggattgcctt ggtaaantcc 1740
cacanctgat aattggcatc act 1763

<210> 71
<211> 2110
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2805526CB1

<400> 71
cagagaacgc gctcgcgtag aggaattccg ggtcctcctc ctcgccctac cttggtttac 60
ctgcagccgc ctagtcctct tctccttctc tttgatgatt cagggcctcg gcttccctca 120
cgatattgca gaaagacaca gctttcctct tcctctccaa accacctcga agcaggggca 180
taattggaat atcatttgga gaaagtgtca tggaagttct gcgtccacag cttataagaa 240
ttgatggccg gaattacagg aagaatccag tccaagaaca gacctatcaa catgaagaag 300
atgaagagga cttctatcaa ggctccatgg agtgtgctga tgagccctgt gatgcctacg 360
aggtggagca gaccccacaa ggattccggt ctactttgag ggcccccagc ttgctctata 420

WEST

```
agcatatagt tggaaagaga ggggacacta ggaagaaaat agaaatggag accaaaactt 480
ctattagcat tcctaaacct ggacaagacg gggaaattgt aatcactggc cagcatcgaa 540
atggtgtaat ttcagcccga acacggattg atgttctttt ggacactttt cgaagaaagc 600
agcccttcac tcacttcctt gccttttcc tcaatgaagt tgaggttcag gaaggattcc 660
tgagattcca ggaggaagta ctggcgaagt gctccatgga tcatggggtt gacagcagca 720
ttttccagaa tcctaaaaag cttcatctaa ctattgggat gttggtgctt ttgagtgagg 780
aagagatcca gcagacatgt gagatgctac agcagtgtaa agaggaattc attaatgata 840
tttctggggg taaaccccta gaagtggaga tggcagggat agaatacatg aatgatgatc 900
ctggcatggt ggatgttctt tacgccaaag tccatatgaa agatggctcc aacaggctac 960
aagaattagt tgatcgagtg ctggaacgtt ttcaggcatc tggactaata gtgaaagagt 1020
ggaatagtgt gaaactgcat gctacagtta tgaatacact attcaggaaa gaccccaatg 1080
ctgaaggcag gtacaatctc tacacagcgg aaggcaaata tatcttcaag gaaagagaat 1140
catttgatgg ccgaaatatt ttaaagttgt ttgagaactt ctactttggc tccctaaagc 1200
tgaattcaat tcacatctct cagaggttca ccgtagacag ctttggaaac tacgcttcct 1260
gtggacaaat tgacttctcc tgaggtggat cttggaaagc actagaaact aaacatcttc 1320
accaggtgct gaagaaaagt gtcttcgttt taattgccaa gcagggatgt ggacatttgg 1380
atggtgactt tcctgggtgg ttccccatag attcaccatt gcctctaatg gtgtctacac 1440
ccgtcatact accagctgag atggtggtgg gcataaggag aatttgtgcc tataaccctt 1500
agtgtgttct ggtttttttt cttttaattt ttaaattgtc gtaaaatact cataaaacat 1560
actgtcttca ccatttttaa gtgcacagtt cagtaacgtt aactgttaat acattcataa 1620
tgctgtgtgg ccgtcaccgc cgtccatctc cataggcttc tcagcttgta aaatggaaac 1680
tgtacccatt aaacagtaat tcccactcct cccagccccc gcagccacca ttctgctttc 1740
tgtctctctg gttttgacta ttctcagtat ctcatataag tggaatcata cagtgacttg 1800
tctttttgtg actggcttat ttcacttagc ataatgtcct caaggttcat ccatgttgtg 1860
tcaggtgaca ggatttcctt cctgtattat acatataatg gaatgttccg ttacatgtgc 1920
gtaccacact ctgttcctcc atcagtgaac acttggcttg cttcctcttg actattggga 1980
gtagtgctaa tacagacacg gggtgcaaa tatctctttg agactctgcc tttaattcct 2040
ttaccaccca gagttttatc tttagcaaaa taaccattaa agttgtttgc cttttaaaaa 2100
aaaaaaaaaa                                                        2110
```

<210> 72
<211> 1493
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2850382CB1

<400> 72

```
tggcggtggc gcggaccacg gcaggagcgg accgggcccg gggctgggcc ggcgtcgaac   60
gcagcgaagg ccggaggatg gaaccaggtg aggagctgga agaggagggc tctccaggtg 120
gccgtgagga tggcttcacc gccgagcacc tggctgcaga ggccatggca gctgacatgg 180
acccctggct agtgtttgat gcccgcacaa cgcctgccac tgagctggat gcctggctgg 240
ccaagtaccc accatcccaa gttacccgct atggggaccc cggttcaccc aactcagagc 300
ctgtgggctg gattgcagtg tatgggcagg gctacagccc caactccggg gacgtgcagg 360
gcctgcaggc agcctgggaa gctctgcaga ccagtgggcg gcccatcaca ccgggtaccc 420
tgcgccagct cgccatcacc caccacgtgc tctcgggcaa gtggcttatg catctggcac 480
cgggcttcaa gctggaccac gcctgggctg gcattgcccg ggccgtggtt gaaggccggc 540
ttcaggtggc caaggtgagc ccacgtgcca aggagggtgg gcgccaggtc atctgtgttt 600
acacggacga cttcacggac cgcttgggtg tactggaggc ggattcagcc atccgtgcag 660
cgggcattaa gtgcctgctc acctacaagc ctgatgtcta cacctacctg ggcatctacc 720
gggccaatcg ctggcacctc tgccccactc tctatgagag ccgtttccag cttgggggta 780
gtgcccgtgg ctcccgagtg cttgaccgtg ccaacaacgt ggaactgacc tagaggggcc 840
aaattgggga gactgcccac tccccctcc atgctggggt tggatcctcc tgtcttcctc 900
cttgtcccat gaaggccaca ccccccagct ctggggactc ctaggtcact tgggaactac 960
ctgcatcttc agtcccсttg aacttctgcc ctctgttcag ggctgacaca agccccacag 1020
gctgggggc tccggttccc tgagggatga gccttcagcc tccctttgta atgctgctcc 1080
tctccactgc ccagcaccat gagttgggtg cagacaccta gaaggagaga cttcttggaa 1140
cgctcatccc ccgctatacc tccccttcct cctgcatctc cccttctttc cttcccccctc 1200
aggagagaga aaacttagtg cttccagccc ttcttggagc cttcatggtc caggggtagg 1260
ggccccactg gcctgagcat gccattttga ggggagggta gttgtgccta cttatcccct 1320
ggcagagggg atgccaggac catggacatg aggcttgccc atccctgcca acttacacag 1380
cctgtaccac tgtccccсct tccttggcta ctttgacatg tgcctgctcc tggcatttca 1440
ataaaacccg gcttgggtct gtgtcttgag tgttttttaa aaaaaaaaa aaa        1493
```

<210> 73

<211> 2930
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2929276CB1

<400> 73
tggagagcgg ggagttgtgt ccaccttgcc gacgtcgcta gccgtggggc tgtcctggga 60
aggcggacgg cgagcgcccg gtgtccgcac tcggccgcct gccgtgcccg tctgcgcccg 120
tgtcatcctc actcgggacg cagggaccgt ttttaaatca caggggcgtg tgtcagcctg 180
ccctaggact tcatgtctat atatttcccc attcactgcc ccgactatct gagatcggcc 240
aagatgactg aggtgatgat gaacacccag cccatggagg agatcggcct cagccccccgc 300
aaggatggcc tttcctacca gatcttccca gacccgtcag attttgaccg ctgctgcaaa 360
ctgaaggacc gtctgccctc catagtggtg gaacccacag aaggggaggt ggagagcggg 420
gagctccggt ggccccctga ggagttcctg gtccaggagg atgagcaaga taactgcgaa 480
gagacagcga aagaaaataa agagcagtag agtccctgtg gactccccatg ggtcatacca 540
gccagcatct gttcctgaac tgtgtttttc ccatcatgac ggaagaagag agtgagccgc 600
aattgttctg aaaatgtcaa acgaggcttc tgttttgcac ctgcagatca ccgagttggt 660
tttcttttct tttcttgcct tttttttttt gaaatttgcc gagcagtgga gccctctgac 720
aatttgcaag gccctctgag aaaggaagct gcttagagcc agggggttag tgggtgaggg 780
gagcgagtgc tgtttttgag atcattatct gaactcaggc agcctagtag aggcagtggt 840
gggattccaa tgggtcttgg tgggtgggag gtggggcatg tgcaaagcaa gcaaggaaca 900
tttggggtaa gaaaacaaac atgaggcaaa agaaaaaata catgttttta agaaaacatt 960
gagcagagaa ctgcagccag gatgcgctca gcagacattc actctggctg ctgggacatc 1020
agaaaacaaa gtcttcatct ctctctccag tttcacccac cccacccttt gctttcattt 1080
caggtgtgtt ggtctatatg acagggagga gagtaaagga gagcaggagc aattggctgc 1140
ctgcaaagcc agctggaggt gaagtgcagg aaaggaaagg tcaccccatt ctactccatg 1200
gcctctctgc tcccagctgt ggtaggctca catagccagt gtgatcggtt tttaagaggc 1260
agtgcttttc agcttttctc cctgatatat ccatttttgct tcccagcact ttttaggagt 1320
agtgagagca cttcctgccc ttgttggaag ccccagggtg gacactcagc acgaaggtct 1380
ctcccttaac tgctgccctt ccaagacttg ctcccgagat ggagtgggcg tggtcttcca 1440
ggctggccct tccttctcct caccgccacc ttccctgccc cagccccagc agccatgggt 1500
acatgggtcc ccagctcacc tatggattcc cgccagtctg cccagctgca gtactcacgc 1560
cccatggggg atcttggtct gtttttcttg tgggagccta gtggagagca gacgtggctt 1620
tttatgtgtc ttgttgggga ggtgacttgc atggtgggga caaggctgtc gtggcaacct 1680
tgggatcgag tttgagacta aaggatgtca tgagatccct ggcttctccc catgttgttc 1740
ccggacaagg gcagaaggga ggcatggcaa gggacctctg ctgtccttac tcaacagtgg 1800
tcctcatccc tccccacctc ccactgcttc ctgcaagggc accagttgta tgagaaagtt 1860
ggcctttgga cttaggattt cttattgtag ctaagagcca tctgaagcag caggttgcag 1920
gacaaatgct tcagtccgcc gagagcagta ccgtgtggcc aagaggtgga ctcagagcct 1980
tccttgagct aaactcggcc aaccaaggca cgcagcatgt cccctcaggt ctccagtcag 2040
tccaggttga ccctcagttc tggacgtgtg tatatagctg tatttaatac ctcaaggtca 2100
ttgtggctct ggggatgcca gggcaggagg acgagggtgc gctgtggaca cagcagtccg 2160
cggaattccg ttctgggaag ccaatggtcg ccggcacccc ttgcttcctc cctctgttgt 2220
ctgcctgtgt gacacacatc aatggcaata acttcttcca actcctcgca gaagtgggag 2280
aggccggcag cctgcaccga gaggggcttt cctctctctt gctccccgct tcgttctgtt 2340
ttggctgcag agagtggttc atccatactc tcattccctc gcctcccctt gtggacgggg 2400
gtcttgcctt ttcaattcct gtgttttggt gtcttccctt atctgctacc ctgaatcacc 2460
tgtcctggtc ttgctgtgtg atgggaacat gcttgtaaac tgcgtaacaa atctactttg 2520
tgtatgtgtc tgtttatggg ggtggtttat tattttttgct ggtccctaga ccactttgta 2580
tgaccgtttg cagtctgagc aggccagggg ctgacagcta atgtcaggac cctcagcggt 2640
ggagcctgct ggggggaccc agctgctctt ggacaagtgg ctgagctcct atctggcctc 2700
ctcttttttt tttttttcaa gtaatttgtg tgtatttcta actgattgta ttgaaaaaat 2760
tcctagtatt tcagtaaaaa tgcctgttgt gagatgaacc tcctgtaact tctatctgtt 2820
ctttttttgag gctcagggag aaactagcat tttttttttt ccaaactact ttttgtcact 2880
gtgacagttg taaataaagt ttgaaaatgc tttccaaaaa aaaaaaaaaa 2930

<210> 74
<211> 4204
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3033039CB1

<400> 74
atgacaatgg atgctctgtt ggctcgattg aaacttctga atccagatga ccttagagaa 60
gaaatcgtca aagccggatt gaaatgtgga cccattacat caactacaag gttcattttt 120
gagaaaaaat tggctcaggc tttactggag caaggaggaa ggctgtcttc tttctaccac 180
catgaggcag gtgtcacagc tctcagccag gacccacaaa ggattttgaa gccagctgaa 240
gggaacccaa ctgatcaggc tggtttttct gaagacagag attttggtta cagtgtgggc 300
ctgaatcctc cagaggagga agctgtgaca tccaagacct gctcggtgcc ccctagtgac 360
accgacacct acagagctgg agcgactgcg tctaaggagc cgcccctgta ctatggggtg 420
tgtccagtgt atgaggacgt cccagcgaga aatgaaagga tctatgttta tgaaaataaa 480
aaggaagcat tgcaagctgt caagatgatc aaagggtccc gatttaaagc tttttctacc 540
agagaagacg ctgagaaatt tgctagagga atttgtgatt atttcccttc tccaagcaaa 600
acgtccttac cactgtctcc tgtgaaaaca gctccactct ttagcaatga caggttgaaa 660
gatggtttgt gcttgtcgga atcagaaaca gtcaacaaag agcgagcgaa cagttacaaa 720
aatccccgca cgcaggacct caccgccaag cttcggaaag ctgtggagaa gggagaggag 780
gacacctttt ctgaccttat ctggagcaac ccccggtatc tgataggctc aggagacaac 840
cccactatcg tgcaggaagg gtgcaggtac aacgtgatgc atgttgctgc caaagagaac 900
caggcttcca tctgccagct gactctggac gtcctggaga accctgactt catgaggctg 960
atgtaccctg atgacgacga ggccatgctg cagaagcgta tccgttacgt ggtggacctg 1020
tacctcaaca cccccgacaa gatgggctat gacacaccgt tgcattttgc ttgtaagttt 1080
ggaaatgcag atgtagtcaa cgtgctttcg tcacaccatt tgattgtaaa aaactcaagg 1140
aataaatatg ataaaacacc tgaagatgta atttgtgaaa gaagcaaaaa taaatctgtg 1200
gaactgaagg agcggatcag agagtattta aagggccact actacgtgcc cctcctgaga 1260
gcggaagaga cttcttctcc agtcatcggg gagctgtggt ccccagacca gacggctgag 1320
gcctctcacg tcagccgcta tggaggcagc cccagagacc cggtactgac cctgagagcc 1380
ttcgcagggc ccctgagtcc agccaaggca gaagattttc gcaagctctg gaaaactcca 1440
cctcgagaga aagcaggctt ccttcaccac gtcaagaagt cggacccgga aagaggcttt 1500
gagagagtgg gaagggagct agctcatgag ctggggtatc cctgggttga atactgggaa 1560
tttctgggct gttttgttga tctgtcttcc caggaaggcc tgcaaagact agaagaatat 1620
ctcacacagc aggaaatagg caaaaaggct caacaagaaa caggagaacg ggaagcctcc 1680
tgccgagata aagccaccac gtctggcagc aattccattt ccgtgagggc gtttctagat 1740
gaagatgaca tgagcttgga agaaataaaa aatcggcaaa atgcagctcg aaataacagc 1800
ccgcccacag tcggtgcttt tggacatacg aggtgcagcg ccttcccctt ggagcaggag 1860
gcagacctca tagaagccgc cgagccggga ggtccacaca gcagcagaaa tgggctctgc 1920
catcctctga atcacagcag gaccctggcg ggcaagagac caaaggcccc ccgtggggag 1980
gaagcccatc tgccacctgt ctcggatttg actgttgagt ttgataaact gaatttgcaa 2040
aatataggac gtagcgtttc caagacacca gatgaaagta caaaaactaa agatcagatc 2100
ctgacttcaa gaatcaatgc agtagaaaga gacttgttag agccttctcc cgcagaccaa 2160
ctcgggaatg gccacaggag gacagaaagt gaaatgtcag ccaggatcgc taaaatgtcc 2220
ttgagtccca gcagccccag gcacgaggat cagctcgagg tcaccaggga accggccagg 2280
cggctcttcc tttttggaga ggagccatca aaactcgatc aggatgtttt ggccgctctt 2340
gaatgtgcag acgtcgaccc ccatcagttc ccggccgtgc acagatggaa gagtgctgtc 2400
ctgtgctact caccctcgga cagacagagt tggcccagtc ccgcggtgaa aggaaggttc 2460
aagtctcagc tgccagatct cagtggccct cacagctaca gtccggggag aaacagcgtg 2520
gctggaagca accccgcaaa gccaggcctg ggcagtcctg ggcgctacag ccccgtgcac 2580
gggagccagc tccgcaggat ggcgcgcctg gctgagcttg ccgccctgta ggcttggcgc 2640
tgggctctcg gtttgttctt catttttaaa gaaggaaggg tcatatgttt attgctaaac 2700
tgtcaaaaag gaatatattc tgattaaatt attactcctc actttgaggg tgtgagaatt 2760
ttagaagatt taaatgttct atataacact tagatttctg atattttgga agaagttaga 2820
agttaatgaa agcaaactca gttaccaatt ttctggaaaa tatccatgtg gtaatggtag 2880
actttttagg tggcaatttc taggtctgaa atatagcaga ggaaagggcg ctgaggcagt 2940
tgcaggcagg cagccctgta cttaccctgt actcacctca tccgacagac gctgtggatg 3000
aggaggggct tggcggaggc gtgagcaccg atgtcccttt gataacctgc actcaccaag 3060
atgaactatt tgccgccctg tcttttcctg ggttgggggg tggcatctga tggtggcaga 3120
gtgcctgttg gttcgcccgt gggtctcatg gttcagacag agggaggtgg acggcaggga 3180
tcagggagcc aggagcgcgc ctcagacttg cagcaaccat tgtgatttgg gttgttcgga 3240
atatttaaat tactgatcag aagatgaaag tagcttttct cttgggaagt cttgcagccc 3300
gtgggagtga taccaggagc aacacagagc tcagcagcgg cgccaaggtg ttccctgttt 3360
cctcagcacg tgagccttca ccgcctgctt cattcaggag ccagtgcagc agtaatacag 3420
tctatacatt gttctgtttt caaatttatc ctgaggcttt gttgagcata aatgattata 3480
cgataaaggt atccgttatt ttggaactca tttcagttgg gatctcctgt atgcagagtg 3540
ttgcatttag aggtttgagt cccatcttgg tttcttgccg tgctgactgt agccttcacc 3600
ttgacttgaa tgaaggtctg tggttggaat gtgtgaggag ccgctgaggt gttcaggagg 3660
tgctgcctgg aggtcggttt cttcctgggt gttacgggca actgctcaca cagttgtttc 3720
tctgtgaaca tttccagtgt ttaatccaaa atgaaaaccc accaatgctt ttgctaactt 3780
cagtgccttt tataaatcat ttttaaattt cctgaacttg ctttttgagg atatacaggg 3840
atattaagta gacgcaggat tgtttttgtt tgtaaaaatt ctgaattgaa actttgtttt 3900
aaaaaaaggc ttctttcttt catatgacaa gagataggtc aggaatattg gaatcaagat 3960

ttaaatgtta aaattcgatt ttgttacaca gggtgtgttc atttgttttg tagcagacaa 4020
gatctagatc ccagacagaa acaacacatg ctattctaaa aagccgcatt ttaaaaggca 4080
ccttggttct caaaagaaat cagaatatgg atattcgtag tgatgatctg ttttctctaa 4140
aatcttacca tattgtctgt atatggttgt aaattcaaat ggaaagtaaa acgttttggc 4200
cctg 4204

<210> 75
<211> 2149
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3213216CB1

<400> 75
ctccgcttag tgattttcca ttcactgacc ccgaggaact tcgctccttg gctatgaatc 60
ctgagctgtc tttgctgtaa tacagagttg agcctaatct ctttccccta ttgtgatgcc 120
cctgttacag tagccttgaa tagtcttcct tacctttttta acaagcgtct gagggaattt 180
ttttccttta atagcttggt acatcaagca ggagcctgac tcctaaacca cgctgttcgg 240
ggtatgctga cattgttgtc tggaatataa cctgatttgg aagtgacaag tgactgagga 300
agtacctgta gaccaggaat actgggccag aagaaaaaaa tactgtctag tttagcaaat 360
tgcagaatgg acagcactga atgttggaac ataaaatttt taaaaggttt tggcttgtga 420
tcaaccagga gggaaacatg gttactgctc gccaggaacc tcgcctggtc ctgatttccc 480
tgacctgcga tggtgacacc ctgactctca gtgcagccta cacaaaggac ctactactgc 540
ctatcaaaac gcccaccaca aatgcagtgc acaagtgcag agtgcacggc ctggagatag 600
agggcaggga ctgtggcgag gccgccgccc agtggataac cagcttcctg aagtcacagc 660
cctaccgcct ggtgcacttc gagcctcaca tgcgaccgag acgtcctcat caaatagcag 720
acttgttccg acccaaggac cagattgctt actcagacac cagcccattc ttgatccttt 780
ctgaggcgtc gctggcggat ctcaactcca ggctagagaa gaaagttaaa gcaaccaact 840
tcaggcccaa tattgtaatt tcaggatgcg atgtctatgc agaggattct tgggatgagc 900
ttcttattgg tgacgtggaa ctgaaaaggg tgatggcttg ttccagatgc attttaacca 960
cagtggaccc agacaccggt gtcatgagca ggaaggaacc gctggaaaca ctgaagagtt 1020
atcgccagtg tgacccttca gaacgaaagt tatatggaaa atcaccactc tttgggcagt 1080
attttgtgct ggaaaaccca gggaccatca aagtgggaga ccctgtgtac ctgctgggcc 1140
agtaatggga accgtatgtc ctggaatatt agatgccttt taaaaatgtt ctcaaaaatg 1200
acaacacttg aagcatggtg tttcagaact gagacctcta cattttcttt aaatttgtga 1260
ttttcacatt tttcgtcttt tggacttctg gtgtctcaat gcttcaatgt cccagtgcaa 1320
aaagtaaaga aatatagtct caataactta gtaggacttc agtaagtcac ttaaatgaca 1380
agacaggatt ctgaaaactc cccgtttaac tgattatgga atagttcttt ctcctgcttc 1440
tccgtttatc taccaagagc gcagacttgc atcctgtcac taccactcgt tagagaaaga 1500
gaagaagaga aagaggaaga gtgggtgggc tggaagaata tcctagaatg tgttattgcc 1560
cctgttcatg aggtacgcaa tgaaaattaa attgcacccc aaatatggct ggaatgccac 1620
ttccctttc ttctcaagcc ccgggctagc ttttgaaatg gcataaagac tgaggtgacc 1680
ttcaggaagc actgcagata ttaattttcc atagatctgg atctggccct gctgcttctc 1740
agacagcatt ggatttccta aaggtgctca ggaggatggt tgtgtagtca tggacgaccc 1800
ctggatcctt gccattcccc tcagctaatg acggagtgct ccttctccag ttcggggtga 1860
aaaagttctg aattctgtgg aggagaagaa aagtgattca gtgatttcag atagactact 1920
gaaaaccttt aaaggggggaa aaggaaagca tatgtcagtt gtttaaaacc caatatctat 1980
tttttaactg attgtataac tctaagatct gatgaagtat attttttatt gccattttgt 2040
cctttgatta tattgggaag ttgactaaac ttgaaaaatg tttttaaaac tgtgaataaa 2100
tggaagctac tttgactagt ttcaaaaaac aaacaaaaaa aaaaggaaa 2149

<210> 76
<211> 1348
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3220944CB1

<400> 76
gcagccgccg ctccggactc acgtgtgagt gaggaagaaa acctgaaaaa gaccccaaag 60
aagaagatga aaatggtaac tggagccgta gcgtcggtgc tggaagacga ggccacagac 120
acttctgata gtgaaggaag ctgtggatcg gaaaaggacc acttttattc tgatgatgac 180
gcaatagaag ctgacagtga gggtgatgct gagccctgtg acaaagaaaa tgaaaatgat 240

WEST

ggagaatcaa gtgttgggac taatatgggc tgggcagatg ctatggctaa agtcctcaac 300
aagaaaactc ctgaaagtaa acctactatt ctggtcaaaa ataagaagct ggaaaaggaa 360
aaagaaaagt taaagcaaga aagactagag aaaataaaac agcgtgataa gaggctggag 420
tgggaaatga tgtgcagagt aaagccagat gttgtccaag acaaagagac agagagaaat 480
cttcagagaa ttgcaacaag gggtgtggtg caattattta atgctgttca gaaacatcaa 540
aagaatgttg atgaaaaggt taaggaagct ggaagttcta tgagaaagcg tgctaagttg 600
atatcaactg tttccaagaa agatttcatc agtgttttga gagggatgga tggaagtaca 660
aatgagactg cttcaagcag gaagaaacca aaagccaaac agactgaagt gaaatcagaa 720
gaaggcccag gttggacgat cctacgtgat gatttcatga tgggagcatc tatgaaagac 780
tgggacaagg aaagtgatgg gccagatgac agcagaccag aatctgcaag tgactctgat 840
acataaagca tcataggaaa tacaattgca gtcgttttat tttttctaga aaaatatgtc 900
atcctctgat agttggggaa ttataaggat accatttgta agaaagccaa aagacttttg 960
ccagatttca tatttcccct tttcatgtac actttatata tacttcatta aaattatatt 1020
ttaaaccctt gtataatttt aagcattgtt cctcagaaca tttgtaaaag gatatatttc 1080
tgcttgacca gcgagatgtg cattttgcca ggatcatatt ggtcatgtct attggtgtat 1140
tatttcagta tcaccaatgt tttcagaaat acagtactaa ttcatcatta aactctttga 1200
agttaatatt tttctgcctt ctaacttata gactcaacta tgtatctgta gtttttggga 1260
atggttggtg ttttttgctt tgtgttggga agttattgag aaaacctata taataaaatt 1320
taaaattata gttaaaaaaa aaaaaaaa 1348

<210> 77
<211> 662
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3224631CB1

<400> 77
ggagctgcca gccgggaggc tgcagccgcg ggttgttaca gctgctggag cagcagcggc 60
ccccgctccc gggaaccgtt cccgggccgt tgatcttcgg ccccacacga acagcagaga 120
ggggcagcag gatgaatgtg ggcacagcgc acagcgaggt gaaccccaac acgcgggtga 180
tgaacagccg tggcatctgg ctctcctacg tgctggccat cggtctcctc cacatcgtgc 240
tgctgagcat cccgtttgtg agtgtccctg tcgtctggac cctcaccaac ctcattcaca 300
acatgggcat gtatatcttc ctgcacacgg tgaaggggac accctttgag acccggacca 360
gggcaaggcg aggctgctaa cccactggga gcagatggat tatggggtcc agttcacggc 420
ctctcggaag ttcttgacca tcacacccat cgtgctgtac ttcctcacca gcttctacac 480
taagtacgac cagatccatt ttgtgctcaa caccgtgtca ttgatgagcg tgcttatccc 540
caagctgctc cagctccacg gagtccggat ttttgggatc aataagtact gagagtgcag 600
caccttcaca tgtccagggt ggtaagggag gggtagggta aaaggcatgt gctgcaacac 660
ta 662

<210> 78
<211> 679
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3242839CB1

<400> 78
gcctgctagg atcagcggtg gtggttccgc gatggtaggc ggcggcgggg tcggcggcgg 60
cctcctggag aatgccaacc ccctcatcta ccagcgctct ggggagcggc ctgtgacggc 120
aggcgaggag gacgagcagg ttccccgacag catcgacgca cgcgagatct tcgatctgat 180
tcgctccatc aatgacccgg agcatccact gacgctagag gagttgaacg tagtagagca 240
ggtgcgggtt caggttagcg accccgagag tacagtggct gtggctttca caccaaccat 300
tccgcactgc agcatggcca cccttattgg tctgtccatc aaggtcaagc ttctgcgctc 360
ccttcctcag cgtttcaaga tggacgtgca cattactccg gggacccatg cctcagagca 420
tgcagtgaac aagcaacttg cagataagga gcgggtggca gctgccctgg agaacaccca 480
cctcttggag gttgtgaatc agtgcctgtc agcccgctcc tgagcctggc ctttgacccc 540
tcagcctgca tactggtatc ctggtcccag ctcctgccag ggctgttacc gttgttttct 600
tgaatcactc acaatgagaa actaacattt tgctttttgt aataaagtta atttatattc 660
agttaaaaaa aaaaaaaaa 679

<210> 79

<211> 1160
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3274451CB1

<400> 79
```
cggaatgttg ccctggtcgc cgccgcggtg gggcgcgggc ccggaggagg cgcccgaggg   60
accaggcctt ccggcgggca aggaccaggg cgagccctgg agcgccggag cggacgcccc 120
cagggagctg ggacaccccc catgggacag gtgaggggct cggacggagc cgtgggggga 180
cagcgactgg gggcctccat gccttcacca cgcgctctac ggaggcttct ccattgtccc 240
tgctggagat tgggtccctc ggtctccctt gcaccaggca cccttgggct ccaggggagg 300
gtgcgtgggg gcggtgtggc tgcagcctca gtccctgccc ctccagccca cagcccccat 360
cccgtagccc ttcctggatc tggagagagc agctgctgct gccccaccac ctgccaggcc 420
cggccactga gcagcccccc ggagccccag atccccgttc cccccaggtg gcctgggagg 480
tggccccctc gaggatgact ccactagcgc cctgggaccc caagtatgaa gccaaagcag 540
gacctcggcc ggtgtggggg gccaactgta gctcaggagc ctcgttctca ggccggacgc 600
tgtgtcaccc ctcattctgg ccgctgtatg aagcagcctc gggcaggggt ctcaggcccg 660
tggcccctgc cacagggcac tggaatggac agcaggcgcc cccagatgca gggttcccgg 720
tggtgtgctg tgaagatgtc ttcctctcgg accctctgct gccccggggg cagcgtgttc 780
ccctgtacct gtccaaggcc ccccagcaga tgatgggctc cctgaaactg ctgccgccgc 840
cccccatcat gtctgccagg gtgctccccc gcccatcacc ctcccggggc ccctccactg 900
cctggctcag cgggccggag ctgatcgctc tcactggcct gctgcagatg agccaggggg 960
agcctaggcc cagctcctcc gcggttggcc ccccagacca tacctctgac ccacccagcc 1020
cctgtggtag ccccagcagt tctcagggtg ctgacctctc tctcccacag accccagaca 1080
cccattgtcc atagccttct cagggcagag tgggctggtt gtgttgacaa taaaacagtg 1140
ttggtttgca aaaaaaaaaa                                             1160
```

<210> 80
<211> 949
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3284256CB1

<400> 80
```
gcgcgaacca tggccggcat ggtggacttc caggatgagg agcaggtcaa gtcctttttg   60
gagaacatgg aggtggagtg caactaccac tgctaccacg agaaggaccc ggacggttgc 120
tatcggctgg tggactattt ggaagggatc cggaagaatt ttgatgaggc tgccaaggtg 180
ttgaagttta actgtgaaga gaaccagcac agtgatagct gctacaaact gggggcctac 240
tatgtgactg gaaaaggtgg tctgacccag gacctgaaag ctgccgccag gtgctttttg 300
atggcgtgtg agaagcctgg aaagaagtca atagcagcat gtcacaacgt tggcctcctg 360
gcacatgatg gacaggttaa tgaggatggc cagcctgact tgggaaaggc cagggactac 420
tacacaaggg cctgtgatgg tggctatact tccagttgct tcaacctcag tgccatgttc 480
ctgcagggtg ccccaggctt tcccaaggac atggacctgg catgtaaata ctccatgaaa 540
gcctgtgacc tgggtcatat ctgggcctgt gccaatgcca gtcgcatgta caagctgggg 600
gatggtgttg ataaggatga ggccaaggcc gaggtgctaa aaaatcgagc ccagcagcta 660
cacaaagaac agcagaaagg tgtccaaccc ttaacatttg ggtaatgtgg tccaccctcc 720
ctcccaggca acaaactgct tgaggctggc agctcctgtt tctgaagact gatgcagccc 780
ttgaaggtca acctgctgga gcaaaaaaac ttgggacttg aatttgtagc tccatttaca 840
tggatccatt gccccagcta ctggagtata gcctacaatg tttatttcag tcaatattcc 900
tttatctggg tgttctgtac aatgtgagat tacagtcaat atgccgttg             949
```

<210> 81
<211> 1517
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3507004CB1

<400> 81

WEST

cgctcccctc tcacgcagcc aacatggctc cagtggagca cgttgtggcg gatgctgggg 60
ctttcctgcg gcatgcggct ctgcaggaca tcgggaagaa catttacacc atccgggagg 120
tggtcactga gattcgggac aaggccacac gcaggcggct cgctgtcctg ccctacgagc 180
tgcggttcaa ggagccctta ccggaatacg tgcggctggt gactgagttt tcaaagaaaa 240
caggagacta ccccagcctc tctgccacgg acatccaagt gcttgcactc acataccagt 300
tggaagcaga gtttgttggg gtgtctcacc taaaacaaga accacagaag gttaaggtga 360
gctcatcgat tcagcaccca gaaacacctc tgcacatttc tggtttccat ctgccctaca 420
agcctaaacc cccacaagaa acagaaaaag gacactcagc ttgtgagcct gagaacctgg 480
aatttagttc cttcatgttc tggagaaacc ctttgcccaa catcgatcat gaactgcagg 540
agctgctgat tgacagaggt gaggacgttc caagtgagga ggaggaggag gaagaaaacg 600
ggtttgaaga cagaaaagat gacagcgatg acgacggggg tggctggata acccccagta 660
acatcaagca gatccagcag gagctggagc agtgtgacgt ccccgaggac gtgcgggttg 720
gctgcctgac cacagacttc gccatgcaga atgttctgct gcagatgggg ctgcacgtgc 780
tggcggtgaa cggcatgctg attcgtgagg cccggagcta catcttgcgc tgccatggct 840
gtttcaagac aacgtctgac atgagccgag tgttctgctc acactgtggg aacaagaccc 900
tgaagaaagt gtccgtgacc gtcagcgacg acggcaccct gcacatgcac ttctcccgca 960
accccaaggt gctgaacccc cgcggcctcc ggtactcgct tcccactccc aaagggggca 1020
aatacgccat caacccccat ctcaccgagg atcagcgctt ccctcagctg cgactctccc 1080
aaaaggccag gcagaaaacc aacgtgttcg cccctgacta catcgccggg gtgtcaccct 1140
ttgtcgagaa tgacatctcc agccgctcag ctaccctgca ggtccgggac agcaccttgg 1200
gagctgggcg gagacgctta aatcccaacg cttccagaaa gaagtttgtg aagaaaaggt 1260
gaagagcgag ttcccgcagg caaattggat gggcgtctgg ccgccgtgga gttccggtga 1320
cccatttccc cagccgtgtc gtctccagga ccacccgatg gaaataacag gcgggcttca 1380
cggtgcggct ctgtccgccc atgccccgct gggtctgcag ggaactggac tgtcccatgg 1440
cctgtgagca ccggagcgcc tggctgcctg ccaaggaagt gcaattgcat aaaaacagaa 1500
agaacaacgc cctggag 1517

<210> 82
<211> 1991
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3585823CB1

<400> 82
cagaaattga gtttccacca tccttagatc tacaccagga gattccccag caggaagatc 60
agaaaaagga agtccccaag aggatactgg accagaactt tggggagccc catataacct 120
ctaggctgcc tccactccca ctgcatattc gaatccagca ggccctcacc agcccacttc 180
ccatgactcc tattctggag ggttctcaca gagctcattc gttgcttttt gaaaacagtg 240
acagcttttc tgaggacagc agtacgctgg gtcggaccag gtctcttccc atcactattg 300
aaatgctaaa agttccagac gatgaagaag aagaggagca aacctgtcca tccacattca 360
gtgaagaaat gacacctacc tcagtcattc ctaaattacc acagtgtcta cgggaggaag 420
aagagaagga gagcgactct gattcagaag gtcccattca gtaccgagat gaagaagatg 480
aagatgaaag ctatcagagt gctctcgcca acaaagtgaa gaggaaagac acactggcaa 540
tgaagttgaa ccacagaccc agtgaaccag agttgaacct gaattcttgg ccttgtaaaa 600
gcaaggagga gtggaatgaa atacggcacc agattggaaa cacactgatc cggcgactga 660
gtcaaagacc aacaccagaa gaactagaac aacgcaatat attgcaacct aaaaatgaag 720
ctgatcgtca ggcagaaaaa cgagaaatta aacgtcggct cactagaaag ctcagtcaaa 780
ggccaactgt cgctgaactc cttgccagga agattctgag gtttaatgaa tatgtagagg 840
taacagatgc tcaagattat gaccggcgag ccgacaaacc ttggaccaaa ctgacccctg 900
ctgacaaggc tgccataaga aaagaattaa atgaatttaa aagctccgag atggaggttc 960
atgaagagag caaacatttt acacgctacc atcgtccatg atgccaaagg ttgagagagg 1020
aatcaacatg gctgctttgc tgcttccttc tccaaagtga catatggagg gaactttagc 1080
acttcccagc acagccagaa ttgcatcctc tgggatcttc tgaggtggac agcactttga 1140
atgtagcatt tcactggaac agagtcttat gtgctgcacc gggggcaaaa caacactttg 1200
tcagtgcttt tgaacctttc aatattgtag catgcttgag gagtttttcc cttactggcc 1260
accaaagttc tgaaccactt gcaggttcca ggttttactg gctgcaccac acccettccc 1320
ctagatgact gcctgtgcag agacacagtt tgcaccatta gccttacctg ccctgccctg 1380
attgtgagac ccaaatgtgt aggctctaaa ttccagccat caaatccaat tcctggtggg 1440
gaaaaccttc tggagacccc caaccttctg ataaaagagt ctctacctcc agggaaagcc 1500
ttcttaccac actggcatat cagatgaaag cattgcactg tacctctcgt aacacagcaa 1560
tacagtcctc ttgaggcact caagcctgag aggaagctca ggatctgaca tgttcttcct 1620
tttcctcaca agtcatcatg attttttatt ttaaaataat ctggaagtaa tggaacttta 1680
gtttttcctg aactccaacc agaatccaaa ttggttagat gaggccaggc gcggtggctc 1740
acgcctgtaa tcccagcact ttgggaggcc gaggtgggtg gatcacctga ggtcaggagt 1800

tcaagaccag cctggccaac atggtgaaac cccatctcta ctaaaaatac aaaaattagc 1860
caggtgtggt ggcgcctggt tgaggcatga gaatcgcttg aacccaggag gtggaggttg 1920
cagtgagcca agatcatgcc tactgcactc cagactggca acaaagtggg actctgtctt 1980
aaaaaaaaaa a 1991

<210> 83
<211> 1238
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3743822CB1

<400> 83
acgctagcgc tgcgatggcg gaggccgtgg agcgcactga cgagctggtc cgggagtacc 60
tgctcttccg cgggttcacg cacacactgc ggcagctgga cgccgagatc aaggcggaca 120
aggagaaggg gttccgggtg gataagattg tggaccagct gcagcagtta atgcaggtgt 180
atgacttggc tgcccttcgg gattattgga gctacttgga gcgtcggctc ttcagccgct 240
tggaggatat atacagaccc acaatccaca agctgaaaac cagcctgttt cgattttatc 300
ttgtctacac aatccagaca aacagaaatg acaaggctca ggagttcttt gcaaagcagg 360
ccacggaact ccagaaccag gctgagtgga aggattggtt tgtcctgccc ttcctgccat 420
ccccggacac caaccccacc tttgctacct acttttctcg acagtgggct gacaccttca 480
ttgtgtccct gcacaacttc ctgagcgtcc tgtttcagtg catgccagtc cctgtgatcc 540
tgaactttga tgcggagtgt cagaggacta accaggttca agaagaaaat gaagttctgc 600
gtcagaagct ttttgcattg caagctgaaa tccaccgact gaagaaagag gagcaacagc 660
cagaagagga agaggccttg gtccaacaca aattgcctcc ttatgtctcc aacatggacc 720
gcctgggggа ctcggaactg tgagtgtgtg tgacggccct ttcctttttct ttgccctctg 780
tgctctcagc cccacagttc cctctgatct tgccttgcct gcagtcctct gtgcctgctg 840
aaggagaatt gcggaagagc ctcctggcac acactgctct tgcttgagag ctttattttt 900
atttttagag atggagtctc gctctgtcgc ccaggctaga gtgccgtggc acaatcttgg 960
ctcaaagcaa cttctgcctc ctgggttcag gagaaccact tgagcccagg aggcggaggt 1020
tgcagtgagc cgagatcacg ccactgggag aacccaactc aaaaaaaaaa aaaaattttt 1080
gggttatgga cctttttgga tttgggacct tggggttaag gcgctcaacc tgtacctttt 1140
ccatttggtg gggtcggag cccgaattcg gaaccatgta aaagctgttt cctggttaaa 1200
attggtaacc ggccaaaatt cccacaaaca ttggggcg 1238

<210> 84
<211> 1448
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3808027CB1

<400> 84
gactacgtct gcatcgaagg atcaggaagt ttgtgctctc tgcgtggcta agtttttcac 60
ctactaggac gggggtgggg tggggagaac aggtgtcctt ctaaaataca gcacaagcta 120
cagcctgcgt ccagccataa cccaggagta acatcagaaa caggtgagaa tgaccacttt 180
aactcaccgg gcccgtcgca ctgaaataag caagaactct gaaaagaaga tggaaagtga 240
ggaagacagt aattgggaga aaagtccaga caatgaagat tctggagact ctaaggatat 300
ccgccttact cttatggaag aagtattgct tctgggacta aaagataaag aggggtacac 360
atctttctgg aatgactgca tatcatcagg cctgcgaggg ggcatcctga tagagctggc 420
catgcggggt cgaatctatc tggaaccccc gaccatgcgt aagaagcgac tactagacag 480
aaaggtactg ctaaagtcag acagcccaac aggtgatgtt ttactggatg aaactctgaa 540
acacatcaaa gcaactgaac ccacagaaac tgtccaaaca tggatagagc tactcactgg 600
tgagacctgg aacccсttca aattacagta ccagctgaga aatgtacgag agcgcatcgc 660
aaagaaccta gtagagaaag gtattctaac cactgagaag cagaatttcc tgctatttga 720
catgactact catccagtga ccaatacaac agagaaacag cgactagtga aaaaacttca 780
agatagtgta ctagagcggt gggtaaatga ccctcagcgt atggacaagc gaacactagc 840
actcctggtg ctagcccact cctctgatgt gctagagaat gtcttctcct ctctgacaga 900
tgacaagtat gatgtggcaa tgaatcgagc caaggactta gtagaactgg accctgaagt 960
ggaagggaca aagcctagtg ccacagaaat gatctgggct gtgctggcag ccttcaataa 1020
atcttaaagc cggcaggtgg atttctcctt ttcccctgct ggctggtgac tgtcagagac 1080
cccatcactg agttttgtgt gatgagatgt tttcatcatt tttgtttcct tttcttgaat 1140
cagacttgtg atttggggag agcaacttca gggcttctcc atagaccttg accattataa 1200

WEST

gcagacctta tttctcccta cacttccact ccataggtaa ataaactggg tgaagccaca 1260
cacacccaat gaacattttc tcccatttgg tggtttagtt tcacccagtg tgcttttgga 1320
taagcagaaa tatctaaatg ggattttgag ttctccagct atagaatgtg aatgagataa 1380
aaagtgaact atctattata tttttctgag tttgttttct atcataaata catgttttca 1440
cagccttt 1448

<210> 85
<211> 807
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4016220CB1

<400> 85
cggacctggg gcctgaccat ggaatcagtt ccattcctgc ctcaatgact tgcaacctgt 60
ttcctcctct gcagattggg aacagcaatc cccacatcaa caatttggcg cccccactct 120
cccaggggaa actgggggac cctctaggcg gtgtttgcag agtcccctcc cagcaggagc 180
cacagggttg aataggagct gccagatcag tctccttccc gggacatccc cgggggcggc 240
agctccacct gctgcaggaa aggtaggacc tgatggctcc aggccctcgg aggctcagag 300
gcgcagggct gagccatggg aacgaagcgc gaggccatcc tgaaggtgct ggagaacctg 360
acaccggagg agctcaagaa gttcaagatg aagctgggga cggtgccgct gcgcgagggc 420
tttgagcgca tcccgcgggg cgcgctcggg cagctagata tcgtggacct caccgacaag 480
ctggtcgcct cctactacga ggactacgca gccgagctcg tcgtggccgt gctgcgcgac 540
atgcgcatgt tggaggaggc cgcacggctg cagcgggctg cgtgagggcc actctgagct 600
ggaggcgggt caaggcaaaa ggagtaatgc aacgcctgtg aagccagccc cacgcgcgca 660
ccagtcgcgt aagacaacag cagtgtctcc acctcgggga ccaaggacgc ctccgtgcct 720
ccagaccccg cctcctccag cccctgcacc tgtcatttat tcttccactg cccaataaat 780
attcatggca gactttaaaa aaaaaaa 807

<210> 86
<211> 1071
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4093555CB1

<400> 86
ggtggggcaa gctgcctaga gaagtgagag cgacgtcagc tattgaccaa tgggaagagc 60
tgatggtatg gcgtgggagc aagagtgaca acgattggtc agccttgcat ctctacgcct 120
aaggcgggag ctcctggagg cggaggccgc gggtgggggc ctagtggagt gaggggtaac 180
aagatggcga ccgagacggt ggagctccat aagctaaagc ttgccgaact aaagcaagaa 240
tgtcttgctc gtggtttgga gaccaaggga ataaagcaag atcttatcca cagactccag 300
gcatatcttg aagaacatgc tgaagaggag gcaaatgaag aagatgtact gggagatgaa 360
acagaggaag aagaaacaaa gcccattgag ctccctgtca aagaggaaga accccctgaa 420
aaaactgttg atgtggcagc agagaagaaa gtggtgaaaa ttacatctga aataccacag 480
actgagagaa tgcagaagag ggctgaacga ttcaatgtac ctgtgagctt ggagagtaag 540
aaagctgctc gggcagctag gtttgggatt tcttcagttc caacaaaagg tctgtcatct 600
gataacaaac ctatggttaa cttggataag ctgaaggaaa gagctcaaag atttggtttg 660
aatgtctctt caatctccag aaagtctgaa gatgatgaga aactgaaaaa gaggaaggag 720
cgatttggga ttgtcacaag ttcagctgga actggaacca cagaggatac agaggcaaag 780
aagaggaaaa gagcagagcg ctttgggatt gcctgatgaa aagttcctga tactttctgt 840
tctccagtgt tttccatttc tctccttctt cttggtcaca tatatgccta aatgcacagt 900
catgtgccta cgtcctgcct cgcaatgagg gagcatgtac cccaggtaca tccatgaact 960
gcggcagcag tttgacttat tgctgtttca gctttaaggt tgttgtgttt ttgtttttga 1020
ttatgttgct tgttaataaa aaaaaataga aaagagaaaa aaaaaaaaaa a 1071

<210> 87
<211> 988
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature

<223> Incyte ID No: 4829366CB1

<400> 87
tcggtagtgc ggcgctgttt aaagatggcg gcggaggaac ctcagcagca gaagcaggag 60
ccgctgggca gcgactccga aggtgttaac tgtctggcct atgatgaagc catcatggct 120
cagcaggacc gaattcagca agagattgct gtgcagaacc ctctggtgtc agagcggctg 180
gagctctcgg tcctatacaa ggagtatgct gaagatgaca acatctatca acagaagatc 240
aaggacctcc acaaaaagta ctcgtacatc cgcaagacca ggcctgacgg caactgtttc 300
tatcgggctt tcggattctc ccacttggag gcactgctgg atgacagcaa ggagttgcag 360
cggttcaagg ctgtgtctgc caagagcaag gaagacctgg tgtcccaggg cttcactgaa 420
ttcacaattg aggatttcca caacacgttc atggacctga ttgagcaggt ggagaagcag 480
acctctgtcg ccgacctgct ggcctccttc aatgaccaga gcacctccga ctaccttgtg 540
gtctacctgc ggctgctcac ctcgggctac ctgcagcgcg agagcaagtt cttcgagcac 600
ttcatcgagg gtggacggac tgtcaaggag ttctgccagc aggaggtgga gcccatgtgc 660
aaggagagcg accacatcca catcattgcg ctggcccagg ccctcagcgt gtccatccag 720
gtggagtaca tggaccgcgg cgagggcggc accaccaatc cgcacatctt ccctgagggc 780
tccgagccca aggtctacct tctctaccgg cctggacact acgatatcct ctacaaatag 840
ggctggctcc agcccgctgc tgccctgctg ccccctctg ccaggcgcta gacatgtaca 900
gaggtttttc tgtggttgta aatggtccta tttcaccccc ttcttcctgt cacatgaccc 960
ccccccatgt tttattaaag gggtgca 988

<210> 88
<211> 1385
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4830720CB1

<400> 88
cccacgcgtc cggccggaag agggagtctg taggggcggg ccggctggcg tccccttcc 60
ggccggtccc catggaggcg ctggggaagc tgaagcagtt cgatgcctac cccaagactt 120
tggaggactt ccgggtcaag acctgcgggg gcgccaccgt gaccattgtc agtggccttc 180
tcatgctgct actgttcctg tccgagctgc agtattacct caccacggag gtgcatcctg 240
agctctacgt ggacaagtcg cggggagata aactgaagat caacatcgat gtactttttc 300
cgcacatgcc ttgtgcctat ctgagtattg atgccatgga tgtggccgga gaacagcagc 360
tggatgtgga acacaacctg ctcaagcaac gactagataa agatggcatc cccgtgagct 420
cagaggctga gcggcatgag cttgggaaag tcgaggtgac ggtgtttgac cctgactccc 480
tggaccctga tcgctgtgag agctgctatg gtgctgaggc agaagatatc aagtgctgta 540
acacctgtga agatgtgcgg gaggcatatc gccgtagagg ctgggccttc aagaacccag 600
atactattga gcagtgccgg cgagagggct tcagccagaa gatgcaggag cagaagaatg 660
aaggctgcca ggtgtatggc ttcttggaag tcaataaggt ggccggaaac ttccactttg 720
cccctgggaa gagcttccag cagtcccatg tgcacgtcca tgacttgcag agctttggcc 780
ttgacaacat caacatgacc cactacatcc agcacctgtc atttggggag gactatccag 840
gcattgtgaa ccccctggac cacaccaatg tcactgcgcc ccaagcctcc atgatgttcc 900
agtactttgt gaaggtggtg cccactgtgt acatgaaggt ggacggagag gctcccctcc 960
caccccaggt actgaggaca aatcagttct ctgtgaccag acatgagaag gttgccaatg 1020
ggctgttggg cgaccaaggc cttcccggag tcttcgtcct ctatgagctc tcgcccatga 1080
tggtgaagct gacggagaag cacaggtcct tcacccactt cctgacaggt gtgtgcgcca 1140
tcattggggg catgttcaca gtggctggac tcatcgattc gctcatctac cactcagcac 1200
gagccatcca gaagaaaatt gatctaggga agacaacgta gtcaccctcg gtgcttcctc 1260
tgtctcctct ttctccctgg cctgtggttg tcccccagcc tctgccaccc tccacctcct 1320
cggtcagccc cagccccagg ttgataaatc tattgattga ttgtgatagt aaaaaaaaaa 1380
aaaaa 1385

<210> 89
<211> 1524
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 5389730CB1

<400> 89
gccgccacct ctccccagcc caggagaggc tgcggagccg cagccgccca gaccgcgcag 60

```
cgcgggaggc aggttccgca cgaaataaat cagaatgagt tatgcagaaa aacccgatga 120
aatcacgaaa gatgagtgga tggaaaagct caataacttg catgtccaga gagcagacat 180
gaaccgcctc atcatgaact acctggtcac agagggcttt aaggaagcag cggagaagtt 240
tcgaatggaa tctggaatcg aacctagtgt ggatctggaa acacttgatg aacgaatcaa 300
gatccgggag atgatactga aaggtcagat tcaggaggcc atcgccttga tcaacagcct 360
ccacccagag ctcttggaca caaaccggta tctttacttc catttgcagc aacagcattt 420
gatcgagctg atccgccagc gggagacaga ggcggcgctg gagtttgcac agactcagct 480
ggcggagcag ggcgaggaga gccgagagtg cctcacagag atggagcgta ccctggcact 540
gctggccttt gacagtcccg aggagtcgcc cttcggagac ctcctccaca ccatgcagag 600
gcagaaggtg tggagtgaag ttaaccaagc tgtgctagat tatgaaaatc gcgagtcaac 660
acccaaactg gcaaaattac tgaaactact actttgggct cagaacgagc tggaccagaa 720
gaaagtaaaa tatcccaaaa tgacagacct cagcaagggt gtgattgagg agcccaagta 780
gcgcctgcgc ttgcgtggtg gatccaacac cagccctgcg tcgtgggact tgcctcagat 840
cagcctgcga ctgcaagatt cttactgcag tagagaactc tttttctccc ttgtactttt 900
ttttgacctg gcatcttttt atagggaaaa atggcctttg taggcagtgg aaaacttgca 960
aggaaagctg ccgtctcttt ggcagtctga tgcagagcct gcactctggc actcgctgaa 1020
gaatctggaa ggttgcggtt tgctcttcca gtgttcgggg gcctctgggct gctgaaggat 1080
tcggtctacc acggagggct gtgctgttag gctgcatccc actcaaaata caggaaaagc 1140
acgaatcatg attctgcttt ctgttagctt aggcagacat tgggccttca cctacaagtt 1200
tttccttacc cctgtggttt ttgtgttttt tttttttct ttttccatag gaaagaatat 1260
ataaatttgt aaatcctaat tcaaagatgg ctcatgtgtg agggcattga gtttgatttg 1320
ttttcccttt ggtctgggtt gtgtggcttt tgggggatgc gtgtgagggg gctatgtgtt 1380
ttttaatttt ttaaatatat attttggtgc tgtgtgtggt aagagacttg ttcctagtgg 1440
atcaatgaac catctcttct gggcagtttt gttgaaaata aaggtttctc tttgatttca 1500
agaatgacca aaaaaaaaaa aaaa                                        1524
```

<210> 90
<211> 1288
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 5397088CB1

<400> 90

```
ccgcccacgt catggcgccc gaggagaacg cggggaccga actcttgctg cagagtttcg   60
agcgccgctt cctggcggca cgcacactgc gctccttccc ctggcagagc ttagaagcaa 120
agttaagaga ctcatcagat tctgagctgc tgcgggatat tttgcacaag actgtgaagc 180
atcctgtgtg tgtgaagcac ccgccgtccg tcaaatatgc ccggtgcttt ctctcagaac 240
tcatcaaaaa gcacgaggct gtccacacag agcctttgga cgagctgtat gaagcgctgg 300
cggagaccct gatggccaag gagtccaccc agggccaccg gagctatttg ctgccctcgg 360
gaggctcggt cacactctcc gagagcacgg ccatcatctc ctacggtacc acaggcctgg 420
tcacatggga cgccgccctc taccttgcag aatgggccat cgagaacccg gcagtcttca 480
ctaacaggac tgtcctagag cttggcagtg gtgctggcct cacaggcctg gccatctgca 540
agatgtgccg cccccgggca tacatcttca gcgactgtca cagccgggct cttgagcagc 600
tccgagggaa tgtccttctc aatggcctct cattagaggc agacatcact gccaagttag 660
acagccccag ggtgacagtg gcccagctgg actgggacgt cgcgacggtc catcagctct 720
ctgccttcca gccagatgtt gtcattgcag cagatgtgct gtattgccca gaagccatca 780
tgtcgctggt cggcgtcctg cggaggctgg ctgcctgccg ggaggaccag cgggctcctg 840
aggtctacgt ggcctttacc gtccgcaacc cagagacgtg ccagctgttc accaccgagc 900
taggccgggc cgggatcaga tgggaagtgg aacctcgtca tgagcagaaa ctgtttccct 960
acgaagagca cttggagatg gcaatgctga atctcaccct gtaggactca cacacgactc 1020
caacgggatt gtgagaatca agtcactctc atgggaagaa tttttatatg ggaaagcgga 1080
taaaactttc attggactgg aatgtttgga gaatgttaat ttccaaatca ggaaccacaa 1140
actgccctct aataagacat cggctatcta agcgtgtggg tgcccccttt ctgccagcag 1200
ttgtggttct taagaaaatc accgataaat cagacatgaa aattctggct ccaaaaacag 1260
cattgtctgt gtgcgagtgg gagcgtgt                                    1288
```

<210> 91
<211> 1161
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 5425521CB1

<400> 91
gccggtggag cgcgagtagg aagtggtgag ttcggagtag agatggccgc gcttgcaccg 60
ctgccccccgc tccccgcaca gttcaagagc atacagcatc atctgaggac ggctcaggag 120
catgacaagc gagaccctgt ggtggcttat tactgtcgtt tatacgcaat gcagactgga 180
atgaagatcg atagtaaaac tcctgaatgt cgcaaatttt tatcaaagtt aatggatcag 240
ttagaagctc taaagaagca gttgggtgat aatgaagcta ttactcaaga aatagtgggc 300
tgtgcccatt tggagaatta tgctttgaaa atgtttttgt atgcagacaa tgaagatcgt 360
gctggacgat ttcacaaaaa catgatcaag tccttctata ctgcaagtct tttgatagat 420
gtcataacag tatttggaga actcactgat gaaaatgtga aacacaggaa gtatgccaga 480
tggaaggcaa catacatcca taattgttta aagaatgggg agactcctca agcaggcccct 540
gttggaattg aagaagataa tgatattgaa gaaaatgaag atgctggagc agcctctctg 600
cccactcagc caactcagcc atcatcatct tcaacttatg acccaagcaa catgccatca 660
ggcaactata ctggaataca gattcctccg ggtgcacacg ctccagctaa tacaccagca 720
gaagtgcctc acagcacagg tgtagcaagt aatactatcc aacctactcc acagactata 780
cctgccattg atcccgcact tttcaataca atttcccagg gggatgttcg tctaacccca 840
gaagactttg ctagagctca gaagtactgc aaatatgctg gcagtgcttt gcagtatgaa 900
gatgtaagca ctgctgtcca gaatctacaa aaggctctca agttactgac gacaggcaga 960
gaatgaagcc tttgtatgac agacccatgt atttttggca tgaggaacta acagtccatt 1020
actctatctt cagcctatca ggatcacagt tttaaggaag acttggtttt gttgaatatg 1080
acaatgaaat ctgtgtgtat cagattttta ttgaagcatt catcagcagc ctcagccagt 1140
tttcattgtc catttactag a 1161

<210> 92
<211> 614
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 5495427CB1

<400> 92
gccgcgacca tctcaagtgt ttcgccacac tcatcatggc gtgggcagta agtcacttcc 60
gcccgggacc ggaagtgtgg gatactgcga gtatggcggc gtcaaaggtg aagcaggaca 120
tgcctccgcc ggggggctat gggcccatcg actacaaacg gaacttgccg cgtcgaggac 180
tgtcgggcta cagcatgctg gccataggga ttggaaccct gatctacggg cactggagca 240
taatgaagtg gaaccgtgag cgcaggcgcc tacaaatcga ggacttcgag gctcgcatcg 300
cgctgttgcc actgttacag gcagaaaccg accggaggac cttgcagatg cttcgggaga 360
acctggagga ggaggccatc atcatgaagg acgtgcccga ctggaaggtg ggggagtctg 420
tgttccacac aacccgctgg gtgccccccct tgatcgggga gctgtacggg ctgcgcacca 480
cagaggaggc tctccatgcc agccacggct tcatgtggta cacgtaggcc ctgtgccctc 540
cggccacctg gatccctgcc cctcccccact gggacggaat aaatgctctg cagacctgga 600
aaaaaaaaaa aaaa 614

<210> 93
<211> 1462
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 003908CB1

<400> 93
gttcggcgca gctaacagac ggcggcagtg cgagaaagcc gaagatggcg gtccccgcgg 60
cgctgatcct acgggagagc cccagcatga agaaagcagt gtcactgata aatgcaatag 120
atacaggaag atttccacgg ttgctcactc ggattcttca aaaacttcac ctgaaggctg 180
agagcagttt cagtgaagaa gaggaagaaa aacttcaagc ggcattttct ctagagaaac 240
aagatcttca cctagttctt gaaacaatat catttatttt agaacaggca gtgtatcaca 300
atgtgaagcc agcagctttg cagcagcaat tagagaacat tcatcttaga caagacaaag 360
ctgaagcatt tgtcaatact tggtcttcta tgggtcaaga aacagttgaa aagttccggc 420
agagaattct ggctccctgt aagctagaga ctgttggatg gcagcttaac cttcagatgg 480
ctcactctgc tcaagcaaaa ctaaaatctc ctcaagctgt gttacaactc ggagtgaaca 540
atgaagattc aaagagcctg gagaaagttc ttgtggaatt cagtcacaag gagttgtttg 600
atttctataa caagctagag actatacaag cacagctgga ttcccttaca tgatgttttc 660
gaagactgtt tttttcatca cgctcctgcc acctcattat tttgcattga agatacattg 720
ccaggttgtg ttttctgaag gattcagtga cttgctttct gtaaattata tggcttatca 780

WEST

cttcttagac aaataacaac caatagagat cattgttaag aatactgagg ttctaatata 840
ctttctttag ttctgtgagc caacagtaat tattaagaac actttccctt taaaggaaac 900
acaagtgaat accatattgt ttttactgtc atagtgttgc tttcttgcct gtcctgctta 960
gtttttactt gctggatgat accataatgt atcaaggagc gtccatggat acaagataag 1020
atgtgtacct tagtagaata cagagctttg gtaattacat gaataaaatt aagaaaatag 1080
ccatatacaa tcaaatacac tatggcattt ttatttgaat atgatgagta tattttgctt 1140
cggaaataat ataggaagga aatgtaaaat agtgagtagt atggtatcag ttaattccag 1200
tctgagcttc tctgtcaact tcagtttctc tctcagttta atgatttaat aatagtccag 1260
gtttttgtgt gtttttcttt atactgcaaa ttaataatga ttcactttat agtttgggag 1320
acagaatcag gtcttgaata aaataattgt aatgagtgct aaatgggcac cattattcga 1380
atcagatacc ttttatattc tctttccata aatacgttga tttctgtcaa taaaattttt 1440
gtgtcttaaa aaaaaaaaaa aa 1462

<210> 94
<211> 1828
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 085915CB1

<400> 94
gcacaaatgg ataataaatc tgaagttcag ttgtggctgt taaagagaat tcaggtaccc 60
attgaagata tacttccttc aaaagaagaa aaaagcaaga ccccacccat gttcctgtgc 120
atcaaagtgg gaaaaccaat gagaaaatcc tttgccactc acactgcagc catggtccag 180
cagtacggca aacggagaaa gcagccagag tactggtttg ctgttcctcg ggagagggtg 240
gatcatttgt acacattctt tgttcagtgg tctcccgatg tctatggaaa agatgccaaa 300
gagcaaggct ttgtggtggt ggagaaggaa gaactgactc atgattgaca acttcttcag 360
tgagccaaca accaagagct gggagatcat cactgttgaa gaggcaaagc gcaggaagag 420
cacatgcagc tactatgaag acgaggacga agaggtgctg cctgtcctac ggccccacag 480
cgcgctcctg gagaatatgc acatcgagca gctggcccga cgccttcctg caagggtgca 540
agggtatcca tggagactgg cctatagcac gttagagcac gggaccagct taaagacgct 600
ctaccggaaa tcggcatcac tagacagtcc tgtcctattg gtcatcaaag atatggataa 660
tcagattttt ggagcatatg caactcatcc tttcaagttc agtgaccact attatggcac 720
aggcgaaact tttctctaca cattcagccc tcattttaag gtctttaagt ggagtggaga 780
aaattcatac tttatcaatg gagacataag ttctttagaa cttggtggtg gagggggacg 840
atttggttta tggctagatg ctgatttata ccacggacga agcaactctt gcagcacttt 900
caataatgat attctttcca aaaaggaaga cttcatagtt caggatctgg aggtgtgggc 960
atttgattga aattcagact gccttaaaat ataacattaa aaagactggg ttcgatcagc 1020
cctcctaaag ctggctggaa aaagaagccc cagcccagcc tgcctcatcc caccccaatg 1080
cttcctttct gccatcatct cagagcatga tcacattgca gaaagattct ggaaggtcca 1140
tgtagagggc agacattgaa gaaagaaact taaaatccag gttgttgaaa agactttgta 1200
ctccccacttc ctccaaatcc atacagtgag gaatcagagt gtttatagat atatgagttg 1260
aatgcaattt ttattttttgg taactgtgaa aaataagata ctgtggatat atacatgcct 1320
tgtgtattta tcagcataat tttcattacc aaaatgtaca gctattattt gccatggaaa 1380
aggttagtct catttagaaa aatcgaaagt gcacagcact taaagggaat atatgaggtt 1440
ttttttttaa aaaaaaaaaa cttttattta ttatttgtag tatattgtct gaaatgtgtc 1500
ggcagttttt tttcttttaa tgtgtcaaat cttgaaatat taaatgtata cattttgtgc 1560
tatgttttgg gaacaaatct gtttgattta tatagtttta tattgaattt tttttgccct 1620
gattgtttag ggtgataggt cttaagcagc atatatctat atatctgtat gtgggtatat 1680
agagatatat gtgtgtgtgt atgtatatgt acatatacat atatatgagt gtataattct 1740
aaatttctaa aaactcatta tgaatgttca tcaatttgac tattataggc cagctttcca 1800
tttagtcaat aaaagcgtac atttttag 1828

<210> 95
<211> 1365
<212> DNA
<213> Homo sapiens

<220>
<221> unsure
<222> 87, 124
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature

<223> Incyte ID No: 478615CB1

<400> 95
aggactgctt tatggacttg catcaggagt agcaacagga ctctggactg tcttcacact 60
aggtggtgtg ggcagcaaag cggctgnctc tccagagctt tcttcccctc tggccaacca 120
gagnctcctg cttctgctgg tgttggccaa tctgacagat gcctcaaatg cgccaaaccc 180
ctacagacaa gccattatgt ccttcaagaa cacacaaggt tttaccaatt cttgagattc 240
tgtatcatgt tgaagaaagg aattcacacc atgtgtatat ggcccttata atattgttga 300
tcccttacgga agatgatggc ttcaacagat ccattcatga agtgatacta aaaaatatta 360
cttggtattc agaacgagtt ttaactgaaa tctccttggg gagtctcctg atcctggtgg 420
taataagaac cattcaatac aacatgacta ggacacgaga caagtacctt cacacaaatt 480
gtttggcagc tttagcaaat atgtcggcac agtttcgttc tctccatcag tatgctgccc 540
agaggatcat cagtttattt tctttgctgt ctaaaaaaca caacaaagtt ctggaacaag 600
ccacacagtc cttgagaggt tcgctgagtt ctaatgatgt tcctctacca gattatgcac 660
aagacctaaa tgtcattgaa gaagtgattc gaatgatgtt agagatcatc aactcctgcc 720
tgacaaattc ccttcaccac aacccaaact tggtatacgc cctgctttac aaacgcgatc 780
tctttgaaca atttcgaact catccttcat ttcaggatat aatgcaaaat attgatctgg 840
tgatctcctt ctttagctca aggttgctgc aagctggagc tgagctgtca gtggaacggg 900
tcctggaaat cattaagcaa ggcgtcgttg cgctgcccaa agacagactg aagaaatttc 960
cagaattgaa attcaaatat gtggaagagg agcagcccga ggagtttttt atcccctatg 1020
tctggtctct tgtctacaac tcagcagtcg gcctgtactg gaatccacag gacatccagc 1080
tgttcaccat ggattccgac tgagggcagg atgctctccc acccggaccc ctccagccaa 1140
gcagcccttc aagttctttt atttctgggt aacagaagta gacagacagg ttacttggtg 1200
tatcttctgt taaagaggat tgcacgagtg tgttttcctc acacactttg atttggagaa 1260
ttggtgctag ttggcaatag ataactcagc gtagatagta ttgcaaaaag gggaggaaat 1320
acacaacaat aataaatgta aaaacctgct attcaaaaaa aaaaa 1365

<210> 96
<211> 1631
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 924880CB1

<400> 96
ggatcaccat tgcggaccaa ggtgaacaac agtcagaaga aaacgcaagt actaagaact 60
tgacaggcct ttccagtggg actgagaaga aacctcatta catggaagtg ctagaaatgc 120
gagccaaaaa cccagtgccc cagctgcgta aatttaaaac caatgtgtta ccttttcgac 180
aaaatgattc atctagtcat tgccagaaga gtgggtctcc tatttcctca gaagagcggc 240
ggcgcaggga taagcagcat cttgatgaca tcacagcagc tcggcttcta ccacttcacc 300
atatgcccac gcagctgctc tccatagaag aatccttggc acttcagaaa cagcagaaac 360
agaattatga ggagatgcaa gcaaagctcg cagcgcaaaa attagctgaa agactgaata 420
ttaaaatgcg gagttataat ccagaagggg agtcttcagg gagataccga gaagtaaggg 480
atgaagatga cgattggtcc tctgatgaat tctgaagata atctcctaaa tcactgacgt 540
tgagatgtca tcatcttaca tcagactttc taactagtat caagatcagt gtcagatatt 600
gttgagggaa gtaattttat aaagttacac aaaggtagtt ataaaaaaag cccagtttgt 660
ctttcagaag gtgactttca tgtgcttgaa aagtttaata tttgaatatt gtgtttaacc 720
acatggtatt aaaatttttgc aatatattgt gtattggtct gatattttag tatatagtag 780
aacatacttt tttttcttta agccaaatga aaagaggtaa ctttgctttt ttcctttttc 840
ttacctatca aatagcattt attacatgtc tttcagtgaa atacttagtt gttccaggca 900
cctaaaatca attaggaaga catagttccc ttcttttttg ggtaatgaag ggagcagtct 960
aaagaattgt atgcatgttt gcatggggtt attgagacat tggatgtgaa atacctagga 1020
agacatgtat ggagaatgct taacaaatgc tcttctcctt tttctgtctt cccttaggaa 1080
gaacgttatc tctcttctga ttaggaagga cttccctttt gagtagatgc tggaaagtgg 1140
aggttttttt tgttggtttg tttgtttggt tttgtgtttt tttttgagat ggggtcttgc 1200
tctgtcgccc aggctggagt gcagtggcag gatctcggct cactgcaatc taggtggagg 1260
tctttttaaa taacctcatg ttacttcagg caaatctggt acaggagaag cccaagctaa 1320
ctgagctggg tgaaggcttg tgcatcacct ggcttgagtt ccttgctgtc acagatgata 1380
ggttcatgta cacaacatta atgtttgata aggaaagcat tttttttttg ttaaaattga 1440
attgtcagta ctttactttt ttccccccaa ataggtatat atagactgca gaagcttcag 1500
tcccatatat aaacgttcat gtcattttag aggattatat ggttgtctcc agagaaatta 1560
acttacatta tcaaaaaatt ctttgtctaa tatattggag tctttaaaag tgttgaagaa 1620
taaaatatta t 1631

<210> 97

<211> 2533
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 955431CB1

<400> 97
gggggagaac gggcggaggc ggtgcgctcg gcgcttcctg ttccggcgcc aggaggagcc 60
gcgcgctgct ggtgctgttg ccgccgctgc tctagctgcc gtcagtcagg ctgcgcccgc 120
gtcttcaggg cccagtccct cggacccatc gccgcttcta gaccctactg cggtctcgga 180
tattgccggg aaaatgtctg atgaattttc gttggcagat gcactacctg aacactcccc 240
tgccaaaacc tctgctgtga gcaatacaaa acctggccaa cctcctcaag gctggccagg 300
ctccaaccct tggaataatc cgagtgctcc atcttcagtg ccatctggac tcccaccaag 360
tgcaacaccc tccactgtgc cttttggacc agcaccaaca ggaatgtatc cctccgtgcc 420
tcccaccgga ccacctccag gacccccagc accctttcct ccttccggac catcatgtcc 480
cccacctggt ggtccttatc cagccccaac tgtgccgggc cctggcccca cagggccata 540
tcctacacca aatatgccct ttccagagct acccagacca tatggtgcac ccacagatcc 600
agctgcagct ggtcctttag gtccatgggg atccatgtct tctggacctt gggcgccagg 660
aatgggaggg cagtatccta cccctaatat gccatatcca tctccaggcc catatcccgc 720
tcctcctcct ccccaagccc ctggggcagc accacctgtt ccatggggca ccgttccacc 780
aggagcctgg ggaccaccag caccatatcc tgcccctaca ggatcgtatc ccacaccagg 840
actctatcct actcccagta atcctttcca agtgccttca ggaccttctg gtgctccacc 900
aatgcctggt ggcccccatt cttaccatta agttaacaat ggacgaagag atgacgcttt 960
gctttttgaa gtacatgtat atgcacatga atgcatatat aaaaattgct ggtttcacta 1020
ttagagggca ttcatgaaag aacaactctt gcacctctca gagaagataa ctgcctcttg 1080
tacttggatg cgtagtacat catatgtata caatcagata aaagcataga agtaaatcat 1140
tcggatgtga ttttatttg gttttcatgg aaagttaaag tgataaagta tattgaatag 1200
ttctttgaca gaatttgttt aaactatgaa actacacact taaaaatcta agatgtggat 1260
tattgttaga atctgcaact tcattggcaa attatttcaa gtattttttct ataatcactt 1320
tccccttcta aataaataaa cttcgagaat aacccatcat aatccaaaca aatgatgcct 1380
caacattttg agctgctctg tcggacaaat aaacctggtc ctcttgaggt tatattttgg 1440
atatacattt ttaaactgtc agtaattatt gtcagatgtg gagttcaata gccagccagt 1500
gttcattttt atccttgagc ttttagtaaa aacttcctgg ttttattttt agtcattggg 1560
tcatacagca ctaaagtctg ctatttatgg aaactaactt ttttgttttt aatccaggcc 1620
aacatgtatg taaattaaat ttttagataa ttgattatct ctttgtacta cttgagattt 1680
gattatgaga tgtgcatatt gctttgggaa gagctcgagg aaggaaataa ttctctcctt 1740
tgttttgaac ctcaaactag ataaacccta ggaattgctt aactgcaaca agtaattttc 1800
attcccacaa aaacctgagg cagctctttt gcccagagcg ttccctgtag ccaccccccac 1860
cccacttgcc cttggttctt tagaaggagc acacacatcc cttgattcct ccctgatgtg 1920
gtaaactggc acactccagg ggtctaaaac ataaaacagt tgtgtttagg gaaccttaag 1980
tcatgcagac atgactgttc tctttgtaca agtgtgaatc aaaatatgta tctctttttc 2040
agagtctggt taagctatgt cattgtctac tgcatagttt cctgagtctg tttgtaaagt 2100
gcttatggct aacagttcag ttctgtattt gttgacaggt aaataagtgg agttgagtgc 2160
catctttgaa aaaattaccc tctagctcta acactgaaaa taataataaa ttgtagatct 2220
ctgcaactaa gtttaaagca gtgtgactgt gttgcttaaa tatcaagtat tgtttataac 2280
caccaaaaaa aaaaaagccc tggtagtttt ttggcacctt atgtttaaat cagattctta 2340
gatttggagt agacctgacc ttgttattta ttagataaca ttttgaatgt atccattgga 2400
tttctaaaat gtattgtgaa tttctcagac aaacaggatt tatgctggag ctctgttttg 2460
cttagaaata aaatatttag tagtttattt ctgctctaat taaaatgtca agaatgccaa 2520
aaaaaaaaaa aaa 2533

<210> 98
<211> 1266
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1275918CB1

<400> 98
tgcgacctga ggagtcacag ttgtgagcaa gttcggtgcc gcgggctggg attagaatta 60
cttgttattg gtaaatagcc actatggaga ctaaggacca gaagaaacaa agaaagaaaa 120
atagtggacc caaagctgca aagaaaaaga agcggcatct gcaggatctc cagctaggag 180
acgaagaaga tgcctggaag agaaatccca aagcttttgc atttcagtct gctgtgtgga 240

WEST

```
tggctcgatc ctttcacagg actcaggatt tgaagacaaa aaagcatcat attccagtgg 300
ttgatcgaac tccactagag cccccaccaa tagtggtagt ggtgatgggg cctccaaaag 360
ttggaaagag cactttgata caatgcctca ttcggaactt caccggcag aagttgacca 420
agatcagagg ccctgtgatg atcgtgtcag gtaaaaagct ccgactcacc attattgaat 480
gtgggtgtga cattaacatg atgattgatc tggctgaagt agcagatctg gtaagtgagc 540
aggggcagcc tggggtgctg atggagactt acagcattgt gataggttat ttaccccgtg 600
atgaagggaa tagagtttta tgattattaa aggaatcatg gtcatcatca aggatacaat 660
agatgtgatt gaattgatga taataataat agtaataaat gttgttatat taataataca 720
aatggaatgt gtacaataaa atgtattcca aaatctaaaa gcaaatcaca ggatagagaa 780
aacctttaat gaacacagct aataagaact cattaacatg catacatata tacctacaaa 840
tatatcatag taagtaccat ttcttcttgg acccttcctg gcactgtgct aactgctttc 900
tacaaattta gcttacataa acccttgata caccccttgag gtgagtaggt attatctgta 960
gtttacataa gatgaaatag agcctcccag cagttaagta acttgtgtga agatgggacc 1020
cttgttccta atggttctag aacctttatc cttagtgata ataagtaagt acatgaattg 1080
cctgaagaag tggtcagagt ttatcataga aaatctagat agtactcagt gtgtggaaaa 1140
atgtacacgc tcttgcattt atggaaatga taatgttcac tatttctgat attctattaa 1200
ctatcttttt tttaaactat cttttaaatt aactaaaata aaatacaaat taccaaaaaa 1260
aaaaaa                                                            1266
```

<210> 99
<211> 1423
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1290896CB1

<400> 99

```
ccggaagtgc ccgcgcaccg gtttccggtg caggtgggga aaatggcggt gtctacagtg 60
ttctcgactt cgtcgctggg aattggtttg tttgttgctg ttggctcagt gtctccatcg 120
gggaaggagg atctggggtt tcctactcca ccatcttcac ctcttgctga ggtcaccacc 180
ttgacatcat atcttgataa gttgtcaact tctctggatt gggagccttt gggggattag 240
ggatcatgct gtctctgtgt gctcaggatt tgacacagat gcttgctctg tcaaggcaca 300
gcctattgtc tcctttgctc agtgtgacat cattcagacg cttctacaga ggtgacagcc 360
caacagattc ccaaaaggac atgattgaaa tccctttgcc tccatggcag gagagaactg 420
atgaatccat agaaaccaaa agagcccgcc tgctctatga gagcagaaag aggggaatgt 480
tggaaaactg cattcttctt agtctttttg ctaaagaaca tctgcagcac atgacagaaa 540
agcagctgaa cctctatgac cgcctgatta acgagcctag taatgactgg gatatttact 600
actgggccac agaagctaaa ccagccccag aaatatttga aaatgaagtc atggccctgc 660
tgagagactt tgctaaaaac aaaaacaaag agcagagact gcgtgcccca gatcttgagt 720
acctctttga aaagccacgt tgagctgtgc tcccacggcct ggcatggggg ttcagtctgt 780
ggatggtaac tacttatgat ggacgttagc cttgcttccg gcttcttaga tgcccagctg 840
ccctacccca gaccactggt cctgcctcaa gtgatggaca taaccctctc ctaggacata 900
catgtaaatg cacaatgtga ctcattctca tactttttttg ttcagctctg agcctcaaaa 960
gtgatttgtc agcagcgctg gcatgcaggc tttgcctggc tgctatctct agaggcaagt 1020
ctgccacgag agggcactgc aggcgaagca gttgtgcttg ctcattgcct cagcccaagt 1080
gtactcaaag aaaagaggca gccagctgtg cgtgctgtac ggaaagcctc ccgccctccc 1140
tgcagctccc cgccctcagt ggcccagggc tttgttgaag tggaactccc cacttccaac 1200
ctggtatggc tccttctgcg aagggaagct gatcccagcc tcctgagctg aatccttcaa 1260
aggcacagca ggcagaatga tggccacagg caggagtggt gtgggcacac tgcttaggag 1320
tcaacatctt catcattggg ctttctttaa aacgtgcact ttgtaatttg aaagagaatt 1380
attaaagcat actgaaaaaa ggaaatttaa aaaaaaaaaa aaa                   1423
```

<210> 100
<211> 911
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1342736CB1

<400> 100

```
ctgtccgagt gtgggcgcct ccgggcctct gcggaccctg ggcagctctg agctcgcgga 60
ggcgtggccg gtgcgcgggg cccgcggcgc gcggggatgg gggtctcggt ggatgtgcac 120
caggtgtaca agtacccctt cgagcaggtg gtcgccagct ttctccgaaa gtaccccaac 180
```


WEST

cccatggata aaaatgtcat ctcagtaaaa atcatggagg aaaaaagaga tgaatcaaca 240
ggggtcatct acagaaagag gattgcaatc tgtcagaacg tggttccaga aattttaagg 300
aaggtgagca ttttgaaagt acctaatatc caattagaag aggagtcatg gctcaatcct 360
cgggaaagaa acatggccat acggagtcac tgccttacgt ggacacagta tgcatccatg 420
aaggaagagt ctgtcttccg ggaaagtatg gaaaacccaa attggacaga gttcattcaa 480
agaggcagga tttcaatcac aggggttgga tttctcaact gtgttttaga aacttttgcc 540
agcacattct tacgacaggg agcccagaag ggaattagaa tcatggagat gctgctaaag 600
gaacagtgtg gtgcccccttt agctgaataa agaatcatca gagagctgta tccatgtcta 660
ctttttattt tatttatttt ttttttgaga atcacttctt ggccacagca catcacagac 720
acagtttctg gaatacaggt ttgaggatac agtgttggca gcttcaagaa gagaagacct 780
tcttgccagg acataaaatg ataccctcct ctgggagcct gcttcgaata gtgggactca 840
gggagataag accttcttgc tggattttta tgacacaatc tctttataat ttaacaaata 900
aaggaaaggg a 911

<210> 101
<211> 1231
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1394209CB1

<400> 101
aatcacaaag aggaaggaag tggcaacaca gctcccgatg atgagaaacc agacacttcc 60
cttattacac aaggtgttcc tactcctggg cccagtgcta atgtagccaa tgatgccatg 120
tcgatattgg aaaccataac ttctcttaac caagaagcca gtgctgctag ggcttcaaca 180
gaaacatcaa atgccaagac cagtgagaga gcgtcaaaaa aacttccatc tcagccaacc 240
actgatacta gtactgacaa agaaagaact tcagaggaca tggctgataa agaaaaatct 300
acagctgact ctggaggtga aggactggaa acagccccaa agtctgaaga gttcagcgac 360
ctcccctgtc cagtcgaaga aattaaaaat tacacaaaag agcataataa tttaattctg 420
ctaaataagg atgttcaaca ggaaagcagt gagcaaaaaa ataaatcaac agacaaaggt 480
gaaaagaagc cagacagcaa tgagaaagga gaaagaaaga aagaaaagaa ggaaaagact 540
gaaaagaaat ttgatcactc aaaaaagagt gaagatacac agaaagttaa agatgaaaaa 600
caagcaaagg aaaaagaagt agagagttta aaacttcctt cagaaaagaa cagtaataaa 660
gctaaaactg ttgaagggac aaaagaagag gaaaataaaa cacagaaaaa aaaatcatct 720
tattataaga acatactcag agctcaatta ctaaattaca cttaaaattt tctgctttaa 780
atacgtgtca aactatcaaa caaggcagag tggtttccag aaaatacaaa ggcaaagtag 840
acttctgtaa aaataactaa actattaatt tataatgtat gatataataa taaaatgaat 900
tttaagtaac tgaagtattt actacctaat aaaaattagg atatccaact atgaactaag 960
ttgcttaaaa taatgtatgt gaataaatta taagacaaat tataaattaa taacttctgt 1020
caatttactc tgtgaacaga tatatgagta cacaaaaata caaatattta ttgctttttt 1080
tcttaagcaa aggggatata agtctccaaa agtattaaat aaacttatta aaatatgaag 1140
gcaaatttga tacttccctg atcccaaaac actctgttct ttaaactttg taggatgttg 1200
ccagttttta tataccctcc tttggcccct c 1231

<210> 102
<211> 1600
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1394647CB1

<400> 102
cccacgcgtc cggattttgt tgtgttcact gctgcatcct aagcacataa aactgcttgg 60
cacataataa gtgctcaaga agcatgaacc gagggagtaa cgggtcagag gcagcccgag 120
gactgagctc tctgagcctg cagagggcgc tccagccccg cccccaaagc taggccccgc 180
ctctcgcagc cggctggcgc actagtgtta gagcctgcgc accaccgttc ttttaactgc 240
gcaggcgcgc cggaagcacc tagagagcgg cgcgtgcgca gcgggagtcg aagcggagat 300
cccggggtcg cgcgagagcc gcaagcggag ttggtgggcg ctatgctatc acccgaggca 360
gagcgagtgc tgcggtacct tgtagaagtg gaggagctcg ccgaggaggt gctggcggac 420
aagcggcaga ttgtggacct ggacactaaa aggaatcaga atcgagaggg cctgagggcc 480
ctgcagaagg atctcagcct ctctgaagat gtgatggttt gcttcgggaa catgtttatc 540
aagatgcctc accctgagac aaaggaaatg attgaaaaag atcaagatca tctggataaa 600
gaaatagaaa aactgcggaa gcaacttaaa gtgaaggtca accgcctttt tgaggcccaa 660

ggcaaaccgg agctgaaggg ttttaacttg aacccccctca accaggatga gcttaaagct 720
ctcaaggtca tcttgaaagg atgagactca agaaccaaga tgggggacca gcaacccccc 780
agggtcatgg aggacccagg accctccaac cttgacacct gtaaggacag gatctgccct 840
gtaaggggcc agccgtcagg aatctggcca tgaaaacctc tttgtagtgc ttggctactc 900
tgtgatggca ggagggaacc ttcagcctgt ctggctgctg gacctggaca ccagggctcg 960
gtggacacaa gatctattga cgggccttgg tagccaccag tgggtgtgtg gggcagtggc 1020
tgtgggggtg taagaatgac tgcaacaggc acttcccaac aatggcctgc tgttcacatg 1080
gaccctgagc aaggaaggag ggagggaggg gcagagtgga gtgtcattcc agcattcctc 1140
tcagaaggga gagaggtttt caggctggtg ccatgcgatt ggaataaagc aggaggctca 1200
tgggtggttg ctgaatgaag aacagaatct tggtgctttg tggctcacca cagccatctg 1260
tggggcaggc acacacacct cccgccagct ccaattttgc actttttccc tgcttgattc 1320
caagagtagg tgctgcctag cagcccttcg tggccactct ttactcagga gggccttgca 1380
gagtcctgca ccaggcctgg gtgagtggat gcgcctctta ccatatgaca cgtgtcaaga 1440
tgcccttccg ccccctctga aagtggggcc cggccagcac tgctcgttac tgtctgcctt 1500
cagtggtctg aggtcccagt atgaactgcc gtgaagtcaa aactcttatg tgttcattaa 1560
gggctcaata aatgttagct gaatgaaaaa aaaaaaaaaa 1600

<210> 103
<211> 2750
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1436854CB1

<400> 103
gccgagtcgg agccggagcc tgagccgcgc gctgtgtctc cgctgcgtcc gccgaggccc 60
ccgagtgtca gggacaaaag cctccgcctg ctcccgcagc cggggctcat ctgccgccgc 120
cgccgcgctg aggagagttc gccgccgtcg ccgcccgtga ggatctgaga gccatgtcgg 180
ccagcagcct cttggagcag agaccaaaag gtcaaggaaa caaagtacaa aatggatctg 240
tacatcaaaa ggatggatta aacgatgatg attttgaacc ttacttgagt ccacaggcaa 300
ggcccaataa tgcatatact gccatgtcag attcctactt acccagttac tacagtccct 360
ccattggctt ctcctattct ttgggtgaag ctgcttggtc tacggggggt gacacagcca 420
tgccctactt aacttcttat ggacagctga gcaacggaga gccccacttc ctaccagatg 480
caatgtttgg gcaaccagga gccctaggta gcactccatt tcttggtcag catggtttta 540
atttctttcc cagtgggatt gacttctcag catggggaaa taacagttct cagggacagt 600
ctactcagag ctctggatat agtagcaatt atgcttatgc acctagctcc ttaggtggag 660
ccatgattga tggacagtca gcttttgcca atgagaccct caataaggct cctggcatga 720
atactataga ccaagggatg gcagcactga agttgggtag cacagaagtt gcaagcaatg 780
ttccaaaagt tgtaggttct gctgttggta gcgggtccat tactagtaac atcgtggctt 840
ccaatagttt gcctccagcc accattgctc ctccaaaacc agcatcttgg gctgatattg 900
ctagcaagcc tgcaaaacag caacctaaac tgaagaccaa gaatggcatt gcagggtcaa 960
gtcttccgcc accccccgata aagcataaca tggatattgg aacttgggat aacaagggtc 1020
ccgttgcaaa agcccccctca caggctttgg ttcagaatat aggtcagcca acccaggggt 1080
ctcctcagcc tgtaggtcag caggctaaca atagcccacc agtggctcag gcatcagtag 1140
ggcaacagac acagccattg cctccacctc caccacagcc tgcccagctt tcagtccagc 1200
aacaggcagc tcagccaacc cgctgggtag cacctcggaa ccgtggcagt gggttcggtc 1260
ataatggggt ggatggtaat ggagtaggac agtctcaggc tggttctgga tctactcctt 1320
cagaacccca cccagtgttg gagaagcttc ggtccattaa taactataac cccaaagatt 1380
ttgactggaa tctgaaacat ggccgggttt tcatcattaa gagctactct gaggacgata 1440
ttcaccgttc cattaagtat aatatttggt gcagcacaga gcatggtaac aagagactgg 1500
atgctgctta tcgttccatg aacgggaaag gccccgttta cttacttttc agtgtcaacg 1560
gcagtggaca cttctgtggc gtggcagaaa tgaaatctgc tgtggactac aacacatgtg 1620
caggtgtgtg gtcccaggac aaatggaagg gtcgttttga tgtcaggtgg atttttgtga 1680
aggacgttcc caatagccaa ctgcgacaca ttcgcctaga gaacaacgag aataaaccag 1740
tgaccaactc tagggacact caggaagtgc ctctggaaaa ggctaagcag gtgttgaaaa 1800
ttatagccag ctacaagcac accacttcca tttttgatga cttctcacac tatgagaaac 1860
gccaagagga agaagaaagt gttaaaaagg aacgtcaagg tcgtgggaaa taaaaggcag 1920
ttctacacag actgcagcaa cggttgcatc tgcatatcct aagaggaaaa aatgaccttc 1980
aagagaatta ggactttttt cttaatttca ctgacttcag agacgattgc agacttgcag 2040
tttaagtatt ggaatttcac aaaagacata ggacttaact ggaaaatgaa aaaaaaaaga 2100
aaaagaaaaa actaaacaaa aaatccctct aggtagttta ggtgaaaaat gtccctttta 2160
ttttggcttt ggttgtgatt tcagagcata atgctatgtt tttttgtctt tttactatgt 2220
ttttcggatt tttaagtccg taagtgcata cagttttctc taatttttaa accctttcct 2280
cctcccattt tgacatttgc acttggagaa cacttgagtt gtgaaggttt tgggcatcca 2340
ccccagaaag tgggaatttg attttatcct tccgaactgg aagaacattt ttatgaagaa 2400

tttttgtcta ggagaatata acagtgttac ccaaggttgt gtctttaagg gtggttcatt 2460
ttctctgacc ttttgttact caaagtaaag tactaggagt cctaagaaat gttctgttct 2520
tgtacattat actgattaag tcaggattaa tttgatttca aagctgagaa cagtggtaaa 2580
aactcgttta cagaaatgca ttttggaaga gaaaaatact gtaaaacgtg tcgtgaatgt 2640
ttcttcagtt tcttgttcag ccaatgagga aagggcattg cctttctttt taccattaat 2700
cacttctcaa taaacgtgag atcctgttga gcatcaaaaa aaaaaaaaaa 2750

<210> 104
<211> 2685
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1447955CB1

<400> 104
ggggcggggc ggaagaaggc ggggagggag ggggcggcac cgcggtcggc tccagtcgcc 60
tccggaggaa ggaagaagag gacgccgggc gcgcagctcg gtgctgatta tcacaactgt 120
ttggtgacct acttcactga cctgtgagga ttccttccct tcaagtactg gattcttgat 180
ctttctgcat catcaagaat ggaaacagac tgtaatccca tggagctaag cagtatgtca 240
ggatttgaag aaggttcaga gctgaacggt tttgaaggaa ctgacatgaa agacatgagg 300
ctcgaagctg aagcagttgt aaatgatgtt ctctttgctg ttaacaacat gtttgtctcg 360
aaaagcctgc ggtgtgcgga tgatgtggcc tatatcaatg tggaaacaaa ggaaagaaac 420
agatattgcc tagaactcac tgaagcaggg ctcaaggtgg taggctatgc ttttgaccag 480
gtagatgatc atttacagac tccctaccat gaaacagtct actccttgtt ggatacactc 540
agccccgcct accgagaagc atttggaaac gcactgcttc aaagactgga agctttgaaa 600
agagatggac agtcatgact acactttttc ctttcagagg ggctggtgct ggtacagaat 660
gttgatataa agcttaaaat tcttgcatat ggtcatagaa aatgcatctt tggttttgtg 720
tttttatcac ttgcttccaa cttaggcttt tggctcagaa gattattgaa taatgatttg 780
tcttagtttc tgtttcagta agggaattct gaggccgttg ctatgatacc atcattaaga 840
cattcacatg tcttcatata atatctcttc atttcaaatc ctaatcacta tttcatacta 900
ttacagggct ttgatgctgc cagcactgtc ttttacatag gaaattctag atttgcacag 960
taatagagga attagaagta cctaactata cactttgatt cagcctgcta aatcaggggt 1020
tcaatactag cttggacaaa ctttgtagta attaattgct accagcctta ttggaaacaa 1080
attatcaact agtttcccct gcacaaattt tgaaattcac tgcttcactt aatctattta 1140
tattactaat aatggattaa taaagatgaa ttaattatat attacttaac tagtattaaa 1200
tgaaaaacag ggactgaaat agttctgtat tccgtgtttg caacagccag ccaactaagc 1260
agaggataaa ccgttagcaa atgaatgtaa taattactca tttccaagat atctaagcac 1320
ataagcaaat acaggaacag acttcattct ttttcttaac aaaaaagcat cttcagtgtg 1380
tgatttaaaa gaaagaagat tctggtttcc tagaaaacaa tattttggcc tgtgttgatt 1440
cttattctga atgtgtgttt acataatgta cagtatatat tcagaaagta tttttgcttc 1500
aacgtttact ttctatgatg tagtgctttg gtattcctac agcaccccac cttccccaac 1560
agatgtacag tgttctgtct ccattcgaaa tctacaatgt aatatgagtg cattgtatgg 1620
gtttgaaacc aaaggatgaa tgaagcattc agagacttaa tatttgaaaa aggaatagtc 1680
agtattttat attttattac aggtactgat atttataaat ttaataaact gtaccatgct 1740
gctgcatgtt ttcaagtaca tgttgaacag taaggattgg ggagttgttt tttaatggtc 1800
acctaaagca gctgctatag aaatgttgaa ctaaaatttt gcatctggtc ataccttcat 1860
gcatttatca tttgcagata tttttccatc attattaaaa aacaggaact tttaggctct 1920
gaagatcatg tggaccagag caaattaaag ttcagtttgt gtcacaattc attgccagac 1980
ttcattggaa tgctttgttt gatgatgtat gttcattctc agctttattt tcagatgctt 2040
aactgggcaa tgaagtctaa cttcaggttg aactttctca tgtttaatct caggctaaat 2100
gtaaatgata tttgtaaagt ttgaataaaa ttctgtttac tcattttgag ttagtatgaa 2160
aaaaagtgat tgtatgttta agaattgaaa ttgttcattt tgtgataaat gattaattcc 2220
aagagcttgt ttccattttt taattctgca ttgttacatt tggatttgaa atgactgaag 2280
tgagctttct cttgttacta ctatgaaaga gcaatgagga acaagaggag aggacattaa 2340
taaaactaca cagatataca tgagaatcta gatattaaaa cctccatttt acctggtctg 2400
tttaaagaac ttaaccagca gataggccag gtgtcgtggc tcacacttgc aattccagtg 2460
gtttgggagg cccaggcagg aggatcactt gaggccagga gtgcaagaca gcctgtgcaa 2520
catagatccc atttctacaa aaaattataa aagcatggcc tggtgacaca tgcctgcagt 2580
cctagctatt aggaaagctg aagcatgaag attgcttgag cccaggaatt ggagactgca 2640
gtgagctaca atcgtgctcg tcctccagcc tggatgcaga cagag 2685

<210> 105
<211> 2439
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1454689CB1

<400> 105
caaagcggga caggtccagc ggttgttctg tgtgaagtcg cgcggcttcc acccacgcag 60
tgttctaagt gaaggccaga aactcgctcg ccatgtcggc tgcagaggcg gggggtgttt 120
tccacagagc caggggcagg accctggccg cgtttcccgc agaaaaggaa agcgaatgga 180
aaggcccatt ctacttcatc ctgggcgcag acccacagtt tgggctgatc aaggcctggt 240
ccactgggga ctgtgacaat ggcggtgacg aatgggaaca ggagatccgt ctaactgagc 300
aagccgtcca ggccatcaac aagctgaacc ccaaacccaa attcttcgtt ctgtgcggcg 360
acctcatcca cgccatgcca gggaagccgt ggcggacgga gcagacggag gacctgaagc 420
gagtgcttag ggcagtggac agggccatcc cactggtcct tgtcagcggc aaccatgaca 480
ttggcaacac cccacggcc gagaccgtcg aggagttctg ccggacttgg ggagatgact 540
acttcagctt ctgggtcggg ggcgtcctgt tcctggtcct caactcccag ttctacgaga 600
acccctccaa atgccccagc ctgaagcagg ctcaggacca gtggctggac gagcagctga 660
gcatcgcgag gcagcggcac tgccagcatg ccatcgtcct ccagcacatc ccgctgttcc 720
tggagagcat cgacgaggac gacgactact acttcaacct cagcaagtcc actcggaaga 780
agttggcaga caagttcatc cacgcaggtg tcagagtcgt gttctcaggc cactaccaca 840
ggaatgccgg gggtacctac cagaacctcg acatggtggt gtcatctgcc attggatgcc 900
agctgggcag agacccccac gggctccgag tcgtggtggt caccgccgag aaaattgttc 960
accgatacta cagtctagat gagctgagtg agaaaggaat agaagacgat ctcatggatt 1020
tgatcaagaa aaaatgacgc tccttcccgt tcccgttcac ttttcacttg cactattttt 1080
ttattttgcc agaaacagca gctgcacaca acctcttgct gaaatataaa aatagcccag 1140
gcaggtttgt gaatttatgt ccttttgcat aaatcagaga gctggggccc tgtcctgaat 1200
tatattcaaa atagaactgc atttccttta aaatggagcg aatgatcctg gaaatcagtt 1260
tttaaaatat tagactatct ctgcatggat gtttgaataa attctcctaa tgctgtgttt 1320
ctcaactcca cttttgaatt ggattgtatt ctggttagca tatggtcaaa gtctgcctat 1380
cgacctcaat tccaatctga cagccaagta tgatctttga taattcctta aacaaactac 1440
acctgattgt attaaattgt ggtcacgatt cccaaaacgc taggcaaaac aaagttcatc 1500
tcctttcctt tccagcggaa cgaacagctg tctcccttgt tgtgagagcc atgcagactt 1560
attttgacaa ggagccaagt aggcaccatt tactgttcaa ttctgagact tctgatgata 1620
ggaacagaat gcagttgata cagctcggag cttaaatttg ccttcctggg tgcctagaaa 1680
agtcaccttt gcacataaat tggccatttc tgaaaatctc ctccaagaga ctgaaaccaa 1740
agctctaaaa cgccgcttct tcagtctctg gggagcacgt tgagttggtt cacatccaca 1800
ttctggctgt gtagccactc cacagatggg tggtctcttg ggatcagaat gcccctccac 1860
tcatgagact cttcattttg tccactttga caggaaaagt gggaatgtat gcagagctct 1920
caaaagaaac aaaaaaggcc aaaacggtgc cttcagccac atcctctgaa ttggccctga 1980
cttggactaa atgcactaat gcaaaatccc ttgacaaaag cgcataggtt atttcaaacc 2040
agcattgttt tttatgtaac ctgttttacc gcatcttctc agcagcttct gaccactgct 2100
caatttctc ctttacagcc attgttctgg tggacaaata acctaggtac tccaaatcct 2160
ggcaggaaaa atatacagca ttatgaaaca gcactcagta atcctaaaat gatttccaaa 2220
gctggttaca catgacctgc aaagtctatt aaattaaaag actttctcat tacagagttt 2280
agtcaacgta gcaaaaccat gggaaattaa atggaaagat aattaaatat gtaaattcat 2340
aaggaacaaa agacgagaag ttaattatac aataggctgg tgacaccaga gttcattgtt 2400
atgtgccaca agaaaattaa attttttaaa aaaaaaaa 2439

<210> 106
<211> 1507
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1568009CB1

<400> 106
caacgcttga gatcctctcc gctcccgcca ccccgcaggt gccccgcgcc gttcccgccg 60
ccccgccgcc cccgtcgcgg gcccctgcac cccgagcatc cgcccccggt ggcacgtccc 120
cgagcccacc aggccggccc cgtctcccca tccgtctagt ccgctcgcgg tgccatgcca 180
ttcctcgggc aggactggcg gtccccccggg cagaactggg tgaagacggc cgacggctgg 240
aagcgcttcc tggatgagaa gagcggcagt ttcgtgagcg acctcagcag ttactgcaac 300
aaggaggtat acaataagga gaatcttttc aacagcctga actatgatgt tgcagccaag 360
aagagaaaga aggacatgct gaatagcaaa accaaaactc agtatttcca ccaagaaaaa 420
tggatctatg ttcacaaagg aagtactaaa gagcgccatg gatattgcac cctgggggaa 480
gctttcaaca gactggactt ctcaactgcc attctggatt ccagaagatt taactacgtg 540
gtccggctgt tggagctgat agcaaagtca cagctcacat ccctgagtgg catcgcccaa 600

WEST

aagaacttca tgaatatttt ggaaaaagtg gtactgaaag tccttgaaga ccagcaaaac 660
attagactaa taagggaact actccagacc ctctacacat ccttatgtac actggtccaa 720
agagtcggca agtctgtgct ggtcgggaac attaacatgt gggtgtatcg gatggagacg 780
attctccact ggcagcagca gctgaacaac attcagatca ccaggcctgc cttcaaaggc 840
ctcaccttca ctgacctgcc tttgtgccta caactgaaca tcatgcagag gctgagcgac 900
gggcgggacc tggtcagcct gggccaggct gcccccgacc tgcacgtgct cagcgaagac 960
cggctgctgt ggaagaaact ctgccagtac cacttctccg agcggcagat ccgcaaacga 1020
ttaattctgt cagacaaagg gcagctggat tggaagaaga tgtatttcaa acttgtccga 1080
tgttacccaa ggaaagagca gtatggagat acccttcagc tctgcaaaca ctgtcacatc 1140
ctttcctgga agggcactga ccatccgtgc actgccaata acccagagag ctgctccgtt 1200
tcactttcac cccaggactt tatcaacttg ttcaagttct gaatcccagc acatgacaac 1260
acttcagaag ggtccccctg ctgactggag agctgggaat atggcatttg gacacttcat 1320
ttgtaaatag tgtacatttt aaacattggc tcgaaacttc agagataagt catggagagg 1380
acattggagg ggagaaatgc agttgctgac tgggaattta agaatgtgaa cttctcacta 1440
gaattggtat ggaaaagcaa aatactgtaa ataaactttt tttctaacaa ttaaaaaaaa 1500
aaaaaaa 1507

<210> 107
<211> 1132
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1677811CB1

<400> 107
tgctgggctt gggtgtccgt cctttcctgt ctccttttct tcctcttttc acttaaatcc 60
actgagtctc tcgccaagtc agtgatttag gtaatcccgc cgcccctggc ggaacctttg 120
tgacgccgga actgatgtgc cgtatggtgg ctgcggtcgg aggactgttt tggtaaagga 180
ctagcagttc tctgcggagg gccggttgat acagttccgg tgggagaacg cggctgcgag 240
gttttcggct ttggctcctg atatgcagcg acagaatttt cggcccccaa ctcctcctta 300
ccctggtccg ggtggaggag gttggggtag cggaagcagc ttccggggaa ccccgggcgg 360
gggcggacca cggccgccct cccctcgaga cgggtacggg agtccgcacc acacgccgcc 420
gtacgggccc cggtctaggc cgtacgggag cagtcactct ccgcgacacg gcggcagctt 480
cccggggggc cggttcgggt ctccgtcccc tggcggctac cctggctcct actccaggtc 540
ccccgcgggg tcccagcagc aattcggcta ctccccaggg cagcagcaga cccacccccca 600
gggttctcca aggacatcta caccatttgg atcagggcgt gttagagaaa aaagaatgtc 660
taatgagttg gaaaattatt tcaagccttc aatgcttgaa gatccttggg ctggcctaga 720
accagtatct gtagtggata taagccaaca atacagcaat actcaaacat tcacaggcaa 780
aaaaggaaga tacttttgtt aacatttctg aaattcaact ggaagcttca tgtgtcagga 840
acatcttgga caaaacttta agttgtgttg atataaattt acccaaagat gatgactttg 900
attggataat tagtaaggtc tttttgttat ttttcatcgt atcaggtatt gttgatatta 960
gagaaaaaag taggataact tgcaacattt agctctggaa gtacctacca cattttagag 1020
atttaccgtt tccatatatt taacattcct ggttacataa tggacatttg tcttttaatg 1080
ttttttcaat gttttaaaat aaaacatttt gtcttctaaa aaaaaaaaaa aa 1132

<210> 108
<211> 1822
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1756951CB1

<400> 108
ggctgtgtag actgttgggt aggctgcgtg ctagcttcgg cgcggatccc tgggcgtccg 60
tacgtcggag tccttcgtcc tccagggtcc ctgttctttg cgccagcggg aaccactatc 120
tctgcactcc tggggttttg ttacatggct gctttcctca aaatgagtgt tagtgtcaat 180
ttcttcagac ctttcaccag gtttttggtg ccatttaccc ttcataggaa gagaaataac 240
ttaacaattt tgcagagata catgtcttcc aaaataccag ctgttactta tcctaaaaat 300
gagagtacac ccccttctga agagctagag ttggataagt ggaaaactac catgaaatct 360
agtgtgcaag aagaatgtgt ttcaacaatc tcaagcagta aggatgaaga tcctctagct 420
gccaccagag agttcattga gatgtggaga ttgcttggca gagaagtacc agaacacatc 480
actgaagaag agctcaaaac ccttatggaa tgtgtttcta acacagcaaa aaaaaaatat 540
ttaaaatatt tatatacgaa ggaaaaagtg aaaaaagcta ggcaaataaa aaaggaaatg 600

aaagcagcag caagggaaga agcaaaaaat atcaagctgc tagaaaccac tgaggaagat 660
aaacagaaaa actttctatt tttacgactt tgggatagga atatggacat agcaatgggc 720
tggaagggtg cccaggccat gcagtttgga caacctttgg tttttgacat ggcttacgaa 780
aattatatga aacgaaaaga attgcagaat actgtttccc agcttttaga aagtgaagga 840
tggaacagaa gaaatgttga tcctttccat atttatttct gcaatctaaa aatagatggt 900
gctttgcaca gagagttagt taaacggtat caagaaaaat gggacaaatt gcttttaaca 960
tcaacagaaa agtctcatgt agatttattt ccaaaggaca gtattatcta tttaactgca 1020
gattctccca atgttatgac tactttcagg catgacaaag tttatgtaat tgggtctttt 1080
gttgataaga gtatgcagcc aggcacatcc ctagccaagg caaaacggct gaacctggca 1140
actgaatgcc ttccattaga taaatattta caatgggaaa ttggtaacaa aaatctcacc 1200
ttagatcaaa tgatacgtat tttgttatgt ctgaaaaaca atggtaattg gcaagaggct 1260
ctgcaattcg ttcccaagag aaaacatact ggttttctgg agatttctca gcattctcaa 1320
gagtttatca acagactaaa gaaggcaaag acttaattca ttttcaaaag gttctctgaa 1380
tgtgcacaga acacgtggct caaatgagaa catttgatgg cttaaaaagt aaatgcgtta 1440
gaaatacagt tctgttaatg tatttcttcc caaacaattc atttttctct tctaaaggta 1500
gtctttccca actgactgta gggttgtgtc ttttcccaat taaatatctg cagaactttg 1560
ggattatact ttgtttactg tagaaagata ataaaaagag ttgtccaaga ttgttgaaca 1620
gaataatctt tatcccagtt aaatagttgt accattggta gacttttta tggaggttcc 1680
tagagggtgg tgccctgggg tgggcttgga agctctgcac cccttccccc atagctttcc 1740
ccgtgcatct ctttgtctgt atgttttgta atatctttta cagtaaactg gtaaatgtgt 1800
ttccttcaaa aaaaaaaaaa aa 1822

<210> 109
<211> 933
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1833547CB1

<400> 109
cacgagcgct tcgccatccg ccgactgccc gccttcacgc tgtcccacct ggagagccac 60
cgtgacggcc agcgcagcag catcatggac gtgcggtccc gggtggattc taagaccctg 120
acccgtaaca cgaggatcat tgcagaggcc ctgactcgag tcatctacaa cctgacagag 180
aaggggacac ccccagacat gccggtgttc acagagcaga tgatccagca ggagcagctg 240
gactcggtga tggactggct caccaaccag ccgcgggccg cgcagctggt ggacaaggac 300
agcaccttcc tcagcacgct ggagcaccac ctgagccgct acctgaagga cgtgaagcag 360
caccacgtca aggctgacaa gcgggaccca gagtttgtct tctacgacca gctgaagcaa 420
gtgatgaatg cgtacagagt caagccggcc gtctttgacc tgctcctggc tgttggcatt 480
gctgcctacc tcggcatggc ctacgtggct gtccagcact tcagcctcct ctacaagacc 540
gtccagaggc tgctcgtgaa ggccaagaca cagtgacaca gccaccccca cagccggagc 600
ccccgccgct ccacagtccc tggggccgag cacgagtgag tggacactgc cccgccgcgg 660
gcggccctgc agggacaggg gccctctccc tccccggcgg tggttggaac actgaattac 720
agagcttttt tctgttgctc tccgagactg gggggggatt gtttcttctt ttccttgtct 780
ttgaacttcc ttggaggaga gcttgggaga cgtcccgggg ccaggctacg gacttgcgga 840
cgagcccccc agtcctggga gccggccgcc ctcggtctgg tgtaagcaca catgcacgat 900
taaagaggag acgccgggaa aaaaaaaaaa aaa 933

<210> 110
<211> 681
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1833723CB1

<400> 110
cgagccgggc cggaagaaga ggtagggcgc cgccgtgaca gattagtcct aaagggaacg 60
gggttgttag ttcaattggc taccggaaaa aaccaggctg ggctgggcgc cgccatgaca 120
accgataccg gaaaaggcgg gtcgttcccc ccggacagcc ctacgccggc aaaggtctcg 180
agatgcaggc ggccctagag gtcaccgctc gctactgtgg ccgggagctg gagcagtatg 240
gccagtgtgt ggcggccaag ccggaatcct ggcagcggga ctgtcactac cttaagatga 300
gcattgccca gtgcacatcc tcccacccaa tcatccgcca gatccgccag gcctgtgctc 360
agccttttga ggccttcgag gagtgtcttc gacagaacga ggcagctgtg ggcaactgtg 420
cagagcatat gcgccgcttc ctgcagtgcg ctgagcaggt gcagccgcca cgctcacctg 480

WEST

```
caactgtgga ggcacagcca cttcctgcct cctgaggact cctctgacgg caggaaaact 540
ggacatgaat gactgccccc ccgcccctcc cctgcagagt ggccagatgg agtcctgagc 600
cctggacatg ggcccggctt tcctggatat caggacttcc aataaataaa gactctgtat 660
actgggaaaa aaaaaaaaaa a                                           681
```

```
LOCUS       AX048112                1507 bp    DNA     linear   PAT 15-DEC-2000
DEFINITION  Sequence 106 from Patent WO0070047.
ACCESSION   AX048112
VERSION     AX048112.1  GI:11876935
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yue,H., Tang,Y.T., Lal,P., Reddy,R., Batra,S., Baughn,M.R.,
            Yang,J., Azimzai,Y., Lu,D.A., Au-Young,J. and Shih,L.L.
  TITLE     Full-length molecules expressed in human tissues
  JOURNAL   Patent: WO 0070047-A 106 23-NOV-2000;

  Query Match             100.0%;  Score 1065;  DB 6;  Length 1507;
  Best Local Similarity   100.0%;  Pred. No. 8e-257;
  Matches 1065;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGCCATTCCTCGGGCAGGACTGGCGGTCCCCCGGGCAGAACTGGGTGAAGACGGCCGAC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        175 ATGCCATTCCTCGGGCAGGACTGGCGGTCCCCCGGGCAGAACTGGGTGAAGACGGCCGAC 234

Qy         61 GGCTGGAAGCGCTTCCTGGATGAGAAGAGCGGCAGTTTCGTGAGCGACCTCAGCAGTTAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        235 GGCTGGAAGCGCTTCCTGGATGAGAAGAGCGGCAGTTTCGTGAGCGACCTCAGCAGTTAC 294

Qy        121 TGCAACAAGGAGGTATACAATAAGGAGAATCTTTTCAACAGCCTGAACTATGATGTTGCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        295 TGCAACAAGGAGGTATACAATAAGGAGAATCTTTTCAACAGCCTGAACTATGATGTTGCA 354

Qy        181 GCCAAGAAGAGAAAGAAGGACATGCTGAATAGCAAAACCAAAACTCAGTATTTCCACCAA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        355 GCCAAGAAGAGAAAGAAGGACATGCTGAATAGCAAAACCAAAACTCAGTATTTCCACCAA 414

Qy        241 GAAAAATGGATCTATGTTCACAAAGGAAGTACTAAAGAGCGCCATGGATATTGCACCCTG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        415 GAAAAATGGATCTATGTTCACAAAGGAAGTACTAAAGAGCGCCATGGATATTGCACCCTG 474

Qy        301 GGGGAAGCTTTCAACAGACTGGACTTCTCAACTGCCATTCTGGATTCCAGAAGATTTAAC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        475 GGGGAAGCTTTCAACAGACTGGACTTCTCAACTGCCATTCTGGATTCCAGAAGATTTAAC 534

Qy        361 TACGTGGTCCGGCTGTTGGAGCTGATAGCAAAGTCACAGCTCACATCCCTGAGTGGCATC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        535 TACGTGGTCCGGCTGTTGGAGCTGATAGCAAAGTCACAGCTCACATCCCTGAGTGGCATC 594

Qy        421 GCCCAAAAGAACTTCATGAATATTTTGGAAAAAGTGGTACTGAAAGTCCTTGAAGACCAG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        595 GCCCAAAAGAACTTCATGAATATTTTGGAAAAAGTGGTACTGAAAGTCCTTGAAGACCAG 654

Qy        481 CAAAACATTAGACTAATAAGGGAACTACTCCAGACCCTCTACACATCCTTATGTACACTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        655 CAAAACATTAGACTAATAAGGGAACTACTCCAGACCCTCTACACATCCTTATGTACACTG 714

Qy        541 GTCCAAAGAGTCGGCAAGTCTGTGCTGGTCGGGAACATTAACATGTGGGTGTATCGGATG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        715 GTCCAAAGAGTCGGCAAGTCTGTGCTGGTCGGGAACATTAACATGTGGGTGTATCGGATG 774
```

```
Qy    601 GAGACGATTCTCCACTGGCAGCAGCAGCTGAACAACATTCAGATCACCAGGCCTGCCTTC 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    775 GAGACGATTCTCCACTGGCAGCAGCAGCTGAACAACATTCAGATCACCAGGCCTGCCTTC 834

Qy    661 AAAGGCCTCACCTTCACTGACCTGCCTTTGTGCCTACAACTGAACATCATGCAGAGGCTG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    835 AAAGGCCTCACCTTCACTGACCTGCCTTTGTGCCTACAACTGAACATCATGCAGAGGCTG 894

Qy    721 AGCGACGGGCGGGACCTGGTCAGCCTGGGCCAGGCTGCCCCCGACCTGCACGTGCTCAGC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    895 AGCGACGGGCGGGACCTGGTCAGCCTGGGCCAGGCTGCCCCCGACCTGCACGTGCTCAGC 954

Qy    781 GAAGACCGGCTGCTGTGGAAGAAACTCTGCCAGTACCACTTCTCCGAGCGGCAGATCCGC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    955 GAAGACCGGCTGCTGTGGAAGAAACTCTGCCAGTACCACTTCTCCGAGCGGCAGATCCGC 1014

Qy    841 AAACGATTAATTCTGTCAGACAAAGGGCAGCTGGATTGGAAGAAGATGTATTTCAAACTT 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1015 AAACGATTAATTCTGTCAGACAAAGGGCAGCTGGATTGGAAGAAGATGTATTTCAAACTT 1074

Qy    901 GTCCGATGTTACCCAAGGAAAGAGCAGTATGGAGATACCCTTCAGCTCTGCAAACACTGT 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1075 GTCCGATGTTACCCAAGGAAAGAGCAGTATGGAGATACCCTTCAGCTCTGCAAACACTGT 1134

Qy    961 CACATCCTTTCCTGGAAGGGCACTGACCATCCGTGCACTGCCAATAACCCAGAGAGCTGC 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1135 CACATCCTTTCCTGGAAGGGCACTGACCATCCGTGCACTGCCAATAACCCAGAGAGCTGC 1194

Qy   1021 TCCGTTTCACTTTCACCCCAGGACTTTATCAACTTGTTCAAGTTC 1065
          |||||||||||||||||||||||||||||||||||||||||||||
Db   1195 TCCGTTTCACTTTCACCCCAGGACTTTATCAACTTGTTCAAGTTC 1239
```

```
LOCUS       AX048112                1507 bp    DNA     linear   PAT 15-DEC-2000
DEFINITION  Sequence 106 from Patent WO0070047.
ACCESSION   AX048112
VERSION     AX048112.1  GI:11876935
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yue,H., Tang,Y.T., Lal,P., Reddy,R., Batra,S., Baughn,M.R.,
            Yang,J., Azimzai,Y., Lu,D.A., Au-Young,J. and Shih,L.L.
  TITLE     Full-length molecules expressed in human tissues
  JOURNAL   Patent: WO 0070047-A 106 23-NOV-2000;
            Incyte Genomics, Inc. (US)

Alignment Scores:
Pred. No.:               1.14e-168      Length:         1507
Score:                   1891.00        Matches:        355
Percent Similarity:      100.00%        Conservative:   0
Best Local Similarity:   100.00%        Mismatches:     0
Query Match:             100.00%        Indels:         0
DB:                      6              Gaps:           0

US-10-061-043A-35 (1-355) x AX048112 (1-1507)

Qy          1 MetProPheLeuGlyGlnAspTrpArgSerProGlyGlnAsnTrpValLysThrAlaAsp 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        175 ATGCCATTCCTCGGGCAGGACTGGCGGTCCCCGGGCAGAACTGGGTGAAGACGGCCGAC 234

Qy         21 GlyTrpLysArgPheLeuAspGluLysSerGlySerPheValSerAspLeuSerSerTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        235 GGCTGGAAGCGCTTCCTGGATGAGAAGAGCGGCAGTTTCGTGAGCGACCTCAGCAGTTAC 294

Qy         41 CysAsnLysGluValTyrAsnLysGluAsnLeuPheAsnSerLeuAsnTyrAspValAla 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        295 TGCAACAAGGAGGTATACAATAAGGAGAATCTTTTCAACAGCCTGAACTATGATGTTGCA 354

Qy         61 AlaLysLysArgLysLysAspMetLeuAsnSerLysThrLysThrGlnTyrPheHisGln 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        355 GCCAAGAAGAGAAAGAAGGACATGCTGAATAGCAAAACCAAAACTCAGTATTTCCACCAA 414

Qy         81 GluLysTrpIleTyrValHisLysGlySerThrLysGluArgHisGlyTyrCysThrLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        415 GAAAAATGGATCTATGTTCACAAAGGAAGTACTAAAGAGCGCCATGGATATTGCACCCTG 474

Qy        101 GlyGluAlaPheAsnArgLeuAspPheSerThrAlaIleLeuAspSerArgArgPheAsn 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        475 GGGGAAGCTTTCAACAGACTGGACTTCTCAACTGCCATTCTGGATTCCAGAAGATTTAAC 534

Qy        121 TyrValValArgLeuLeuGluLeuIleAlaLysSerGlnLeuThrSerLeuSerGlyIle 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        535 TACGTGGTCCGGCTGTTGGAGCTGATAGCAAAGTCACAGCTCACATCCCTGAGTGGCATC 594

Qy        141 AlaGlnLysAsnPheMetAsnIleLeuGluLysValValLeuLysValLeuGluAspGln 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        595 GCCCAAAAGAACTTCATGAATATTTTGGAAAAAGTGGTACTGAAAGTCCTTGAAGACCAG 654
```

```
Qy    161 GlnAsnIleArgLeuIleArgGluLeuLeuGlnThrLeuTyrThrSerLeuCysThrLeu 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    655 CAAAACATTAGACTAATAAGGGAACTACTCCAGACCCTCTACACATCCTTATGTACACTG 714

Qy    181 ValGlnArgValGlyLysSerValLeuValGlyAsnIleAsnMetTrpValTyrArgMet 200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    715 GTCCAAAGAGTCGGCAAGTCTGTGCTGGTCGGGAACATTAACATGTGGGTGTATCGGATG 774

Qy    201 GluThrIleLeuHisTrpGlnGlnGlnLeuAsnAsnIleGlnIleThrArgProAlaPhe 220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    775 GAGACGATTCTCCACTGGCAGCAGCAGCTGAACAACATTCAGATCACCAGGCCTGCCTTC 834

Qy    221 LysGlyLeuThrPheThrAspLeuProLeuCysLeuGlnLeuAsnIleMetGlnArgLeu 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    835 AAAGGCCTCACCTTCACTGACCTGCCTTTGTGCCTACAACTGAACATCATGCAGAGGCTG 894

Qy    241 SerAspGlyArgAspLeuValSerLeuGlyGlnAlaAlaProAspLeuHisValLeuSer 260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    895 AGCGACGGGCGGGACCTGGTCAGCCTGGGCCAGGCTGCCCCCGACCTGCACGTGCTCAGC 954

Qy    261 GluAspArgLeuLeuTrpLysLysLeuCysGlnTyrHisPheSerGluArgGlnIleArg 280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    955 GAAGACCGGCTGCTGTGGAAGAAACTCTGCCAGTACCACTTCTCCGAGCGGCAGATCCGC 1014

Qy    281 LysArgLeuIleLeuSerAspLysGlyGlnLeuAspTrpLysLysMetTyrPheLysLeu 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1015 AAACGATTAATTCTGTCAGACAAAGGGCAGCTGGATTGGAAGAAGATGTATTTCAAACTT 1074

Qy    301 ValArgCysTyrProArgLysGluGlnTyrGlyAspThrLeuGlnLeuCysLysHisCys 320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1075 GTCCGATGTTACCCAAGGAAAGAGCAGTATGGAGATACCCTTCAGCTCTGCAAACACTGT 1134

Qy    321 HisIleLeuSerTrpLysGlyThrAspHisProCysThrAlaAsnAsnProGluSerCys 340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1135 CACATCCTTTCCTGGAAGGGCACTGACCATCCGTGCACTGCCAATAACCCAGAGAGCTGC 1194

Qy    341 SerValSerLeuSerProGlnAspPheIleAsnLeuPheLysPhe 355
          |||||||||||||||||||||||||||||||||||||||||||||
Db   1195 TCCGTTTCACTTTCACCCCAGGACTTTATCAACTTGTTCAAGTTC 1239
```